FERROMAGNETIC ISING MEASURES ON LARGE LOCALLY TREE-LIKE GRAPHS

ANIRBAN BASAK* AND AMIR DEMBO†

ABSTRACT. We consider the ferromagnetic Ising model on a sequence of graphs G_n converging locally weakly to a rooted random tree. Generalizing [32], under an appropriate "continuity" property, we show that the Ising measures on these graphs converge locally weakly to a measure, which is obtained by first picking a random tree, and then the symmetric mixture of Ising measures with $+$ and $-$ boundary conditions on that tree. Under the extra assumptions that G_n, of uniformly bounded degrees, are edge-expanders, and ergodicity of the simple random walk on the limiting tree, we show that the local weak limit of the Ising measures conditioned on positive magnetization, is the Ising measure with $+$ boundary condition on the limiting tree. We confirm the "continuity" and ergodicity properties in case of limiting (multi-type) Galton Watson trees, and the edge-expander property for the corresponding configuration model graphs.

1. INTRODUCTION

An *Ising model* on a finite undirected graph $G = (V, E)$, is the following probability distribution over $\underline{x} = \{x_i \;:\; i \in V\}$ with $x_i \in \{-1, +1\}$,

$$\mu(\underline{x}) = \frac{1}{Z(\beta, B)} \exp\Big\{\beta \sum_{(i,j)\in E} x_i x_j + B \sum_{i\in V} x_i\Big\}. \tag{1.1}$$

These distributions are parametrized by *external magnetic field* B and *inverse temperature* parameter β. When $\beta \geq 0$ the model is said to be *ferromagnetic*, and it is termed to be *anti-ferromagnetic* otherwise. Here $Z(\beta, B)$ is the normalizing constant (also known as *partition function*).

Ising model is a paradigm model in statistical physics [35], with much recent in also the Ising model on *non-lattice complex networks* (see [33], and the references therein). In this paper we focus on sparse graphs with locally tree-like structure that is graph sequences $\{G_n\}_{n\in\mathbb{N}}$ converging locally weakly to (random) trees; for a formal definition see Definition 1.1. Study of statistical physics model on such graphs is motivated by numerous examples from combinatorics, computer science and statistical inference (for details, see [11, 31]). The key to such studies is the asymptotics of log partition function, appropriately scaled, as done for example in [10, 20, 37]. In particular [12] shows that for any sequence of graphs locally weakly converging to random trees, the asymptotic *free entropy density* of the Ising model exists, i.e.,

$$\phi(\beta, B) \equiv \lim_{n\to\infty} \phi_n(\beta, B) := \lim_{n\to\infty} \frac{1}{n} \log Z_n(\beta, B). \tag{1.2}$$

Date: March 4, 2013.

2010 *Mathematics Subject Classification.* 05C05, 05C80, 05C81, 60J80, 82B20, 82B26.

Key words and phrases. Ising model, random sparse graphs, Gibbs measures, local weak convergence.

*Research supported by Melvin and Joan Lane endowed Stanford Graduate Fellowship Fund.

†Research partially supported by NSF grant DMS-1106627.

Part of the work was done when the authors participated in the MSRI program on Random Spatial Processes.

Beyond that, perhaps the most interesting feature of the distribution in (1.1) is its "phase transition" phenomenon. Namely, for a wide class of graphs, the measure $\mu(\cdot)$, for large enough β and $B = 0$ decomposes into convex combination of well-separated simple components. This has been shown for the complete graph [17], and for grids [1, 9, 14, 19].

In the context of tree-like graphs G_n, where the neighborhood of a typical vertex has, for large n, approximately the law of the neighborhood of the root of a randomly chosen limiting tree, this picture is only proven for a k-regular limit, see Montanari, Mossel and Sly [32]. We show here the *universality* of this phenomenon, applicable for a general sequence of locally tree-like graphs, including in particular, Erdös-Rényi graphs, random uniform k-partite graphs, and random graphs of a given degree distribution. More precisely, one expects that the marginal distribution of $\mu_n(\,\cdot\,)$ converges to the marginal distribution on a neighborhood of the root for some Ising Gibbs measure on the limiting tree T. Denoting by $\nu_{\pm,\mathsf{T}}$ the Ising Gibbs measures on T, corresponding to plus and minus boundary conditions, for $B > 0$ it easily follows from [12] that, the limiting measure is given by first picking the random tree T, and then conditioned on T, using the Ising Gibbs measure $\nu_{+,\mathsf{T}}$ (the same applies for $B < 0$ with $\nu_{+,\mathsf{T}}$ is replaced by $\nu_{-,\mathsf{T}}$). Recall that for $B = 0$ and β large, there are uncountably many Ising Gibbs measures, hence the convergence to a particular Gibbs measure is not at all clear, as is the choice of the correct Gibbs measure. As demonstrated in [32], for k-regular trees, the plus/minus boundary conditions play a special role. Indeed, it is shown in [32] that if G_n's converge locally weakly to k-regular trees $\mathsf{T} = \mathsf{T}_k$ then, for any $\beta > 0$ and $B = 0$,

$$\mu_n(\,\cdot\,) \to \frac{1}{2}\nu_{+,\mathsf{T}}(\,\cdot\,) + \frac{1}{2}\nu_{-,\mathsf{T}}(\,\cdot\,). \tag{1.3}$$

It is further shown there that, when the graphs $\{G_n\}_{n\in\mathbb{N}}$ are *edge-expanders* ,

$$\mu_{n,\pm}(\,\cdot\,) \to \nu_{\pm,\mathsf{T}}(\,\cdot\,), \tag{1.4}$$

where $\mu_{n,+}(\,\cdot\,)$ and $\mu_{n,-}(\,\cdot\,)$ are the measures (1.1) conditioned to, respectively, $\sum_{i\in V} x_i > 0$ and $\sum_{i\in V} x_i < 0$. The latter sharp result provides a better understanding of $\mu_n(\,\cdot\,)$, and is much harder to prove than (1.3). For genuinely random limiting trees, one expects (1.3) and (1.4) to apply where now T is chosen according to the limiting tree measure.

Key estimates in the proof of (1.3) and (1.4) in [32], involve explicit calculations which crucially rely on regularity of both graph sequence, and the limiting tree. Several new ideas are necessary in the absence of such regularity. For example, the key to the proof of (1.3) in [32] is the continuity, for k-regular infinite trees, of root magnetization under $\nu_{+,\mathsf{T}_k}(\cdot)$, obtained there out of its representation as the largest zero of a real analytic function. No such representation is known for any other tree measure, and for $\beta > \beta_c$, continuity of root magnetization under $\nu_{+,\mathsf{T}}(\cdot)$ is first shown here[1], for a large class of limiting measures (see Section 5). These include *Multitype Galton Watson* (MGW) trees, which arise as the limit of many natural locally tree-like graph ensembles. The proof of (1.4) relies on choosing functionals $F_l(\cdot)$ of the spin configurations on G_n, which approximate the indicator on the vertices that are in "$-$ state", and whose values concentrate as $n, l \to \infty$. The regularity of the graphs G_n, and that of their limit, provide for such functionals, and allows explicit computations involving them, both of which fail as soon as we move away from the regular regime. At the level of generality of our setting the only tools are *unimodularity* of the limiting tree, and properties of simple random walk on it. Hence, a completely different choice of functionals is required here. With $F_l(\cdot)$ defined via *average occupation measure* of the simple random walk on the tree, we show here that (1.4) holds under the same continuity property, for *any* edge-expander G_n's of bounded degree, provided T is ergodic (see Theorem 1.7). We also confirm that the relevant

[1]for $\beta = \beta_c$ one may use the equivalent capacity criterion provided in [36].

MGW trees are ergodic, and the corresponding configuration models are edge-expanders (see Section 5). Thus, our theorem applies for most naturally appearing locally tree-like graphs. (even in non-ergodic setting we show here that every limit point is either $\frac{1}{2}(\nu_{+,\mathsf{T}}+\nu_{-,\mathsf{T}})$, or $\nu_{+,\mathsf{T}}$)

An interesting byproduct of our results is the continuity of percolation probability for *Random Cluster Model*, with $q = 2$, and *wired boundary condition* (see [22] for details on RCM, and its connection with Ising model). Another interesting byproduct of this work is the uniqueness of the *splitting Gibbs measure* (for a definition see [18, Chapter 12]), for $B = 0$ and any boundary condition strictly larger than the free boundary condition (see Lemma 1.15 and Remark 1.16). Many of the techniques developed here should extend to more general settings, e.g. the Potts model.

1.1. **Graph preliminaries and local weak convergence.** In a connected undirected graph $G = (V, E)$ *distance* between two vertices v_1 and v_2 is defined to be the length of the shortest path between them. For a vertex $v \in V$, $B_v(r)$ will denote a ball of radius r around the vertex v i.e. it is the collection of all vertices in G such that its distance from v is less than equal to r. When $r = 1$, i.e. the set of all adjacent vertices to v will often be denoted by ∂v and $\Delta_v := |\partial v|$.
A *rooted graph* (G, o) will be a graph G with a specified vertex o of G which will be denoted as *root* and a *rooted network* is a rooted graph where the graph has marks on its vertices. A *rooted isomorphism* of rooted graphs (or networks) is an isomorphism which maps root of one to that of another. A *rooted network* $(\overline{G}, o)$ will be a rooted graph (G, o) with marks on its vertices.
Let $\mathcal{G}_*$ be the space of rooted isomorphism classes of rooted *connected* locally finite graphs. Similarly for rooted networks we can define $\overline{\mathcal{G}}_*$ as the space of rooted isomorphism classes of rooted connected locally finite networks. Define a metric on $\mathcal{G}_*$ (and on $\overline{\mathcal{G}_*}$) by letting the distance between (G_1, o_1) and (G_2, o_2) be $1/(\alpha + 1)$ where α is the supremum over $r \in \mathbb{N}$ such that there is a rooted isomorphism of balls of radius r around the roots of G_i (and each pair of corresponding marks has distance less than $1/r$). It can be further shown that $\mathcal{G}_*$ (and $\overline{\mathcal{G}}_*$) is a complete separable metric space under this metric (see [5, 7]). Under this metric topology the Borel σ-algebra on $\mathcal{G}_*$ and $\overline{\mathcal{G}}_*$ will be denoted by $\mathscr{C}_{\mathcal{G}_*}$ and $\mathscr{C}_{\overline{\mathcal{G}}_*}$ respectively. For μ_n, μ, probability measures on $\mathcal{G}_*$ (or $\overline{\mathcal{G}}_*$), we will write $\mu_n \Rightarrow \mu$ when μ_n converges weakly to μ with respect to the metric discussed above. Now we are ready to define *local weak convergence* of graphs.

Definition 1.1. *For a finite graph G let $U(G)$ be the probability distribution on $\mathcal{G}_*$ obtained by choosing a uniform random vertex as a root. Let $\{G_n\}_{n\in\mathbb{N}}$ be a sequence of graphs with vertex set $[n] := \{1, 2, \ldots, n\}$ then we say G_n converges locally weakly to μ if $U(G_n) \Rightarrow \mu$ where μ is a probability measure on $\mathcal{G}_*$.*

Definition 1.2. *The graph sequence $\{G_n\}_{n\in\mathbb{N}}$ is said to be* uniformly sparse *if Δ_{I_n} are uniformly integrable under uniform measure on $[n]$ i.e. if*

$$\lim_{L\to\infty} \limsup_{n\to\infty} \frac{1}{n} \sum_{i\in[n]} \Delta_i \mathbb{I}(\Delta_i \geq L) = 0. \tag{1.5}$$

In our setting the limiting object is a (random) tree, and the graphs will be uniformly sparse, hence when writing $G_n \overset{\text{LWC}}{\Longrightarrow} \mu$ we will implicitly assume that μ supported only on trees, and $\{G_n\}_{n\in\mathbb{N}}$ is uniformly sparse. We use $\mathcal{G}_*$ and $\overline{\mathcal{G}}_*$ even when we restrict ourselves to the space of all rooted connected locally finite trees.
In [7] it is shown that any LWC limit must be *involution invariant* and later in [4] this is improved upon and shown that LWC limits must be *unimodular*. To define this notion first observe that, similarly to the space $\mathcal{G}_*$ one can define $\mathcal{G}_{**}$, the space of all isomorphism classes of locally finite

connected graphs with an ordered pair of distinguished vertices and the natural topology thereon. A function f on $\mathcal{G}_{**}$ will be written as $f(G,x,y)$.

Definition 1.3. *Let μ be a probability measure on $\mathcal{G}_*$. We call μ unimodular, denoted by $\mu \in \mathcal{U}$, if for any Borel function $f\ :\mathcal{G}_{**} \to [0,\infty]$*

$$\int \sum_{x \in V(G)} f(G,o,x) d\mu([G,o]) = \int \sum_{x\in V(G)} f(G,x,o) d\mu([G,o]) \tag{1.6}$$

This property of an LWC limit turns out to be very useful and (1.6) has been used in the proofs on several occasions.

1.2. **Local weak convergence of Ising measures.** Before defining the notion of local weak convergence of Ising measures we note that the space $\{(G,\underline{x}_G),\ G \in \mathcal{G}_*,\ \underline{x}_G \in \{-1,+1\}^G\}$ can be identified with $\overline{\mathcal{G}}_*$ where the mark space is $\{-1,1\}$. Thus any probability measure ν on $\{(G,x_G),\ G\in\mathcal{G}_*,\ x_G \in \{-1,1\}^G\}$, can be viewed as a probability measure on $\overline{\mathcal{G}}_*$. The space of all probability measures on $(\overline{\mathcal{G}}_*, \mathscr{C}_{\overline{\mathcal{G}}_*})$ will be denoted by $\mathcal{P}(\overline{\mathcal{G}}_*)$ and any probability measure on $\mathcal{P}(\overline{\mathcal{G}}_*)$ will be denoted by $\mathfrak{m}$. For any $\mathfrak{m}$ and any $t>0$, $\mathfrak{m}^t$ will denote the probability measure on the space of all probability measures on $(\overline{\mathcal{G}}_*(t), \mathscr{C}_{\overline{\mathcal{G}}_*(t)})$ induced by the usual projection, where $\overline{\mathcal{G}}_*(t)$ is the collection of all rooted connected locally finite networks truncated at height t and $\mathscr{C}_{\overline{\mathcal{G}}_*(t)}$ is the Borel σ-algebra on it. Similarly for $\nu \in \mathcal{P}(\overline{\mathcal{G}}_*)$, ν^t will denote the projection of ν on $(\overline{\mathcal{G}}_*(t), \mathscr{C}_{\overline{\mathcal{G}}_*(t)})$. Adapting the Definition 2.3 of [32] to the case of non-deterministic graph limit we next define,

Definition 1.4. *Consider a sequence of graphs $\{G_n\}_{n\in\mathbb{N}}$ and a sequence of Ising measures $\{\mu_n\}_{n\in\mathbb{N}}$ on them. Let $\mathbb{P}_n^t(i)$ denote the law of the pair $(B_i(t), \underline{x}_{B_i(t)})$ when $\underline{x}$ drawn according to μ_n and $i\in[n]$ is a vertex in the graph. Let U_n denote the uniform measure over a random vertex $I_n \in [n]$, and hence $\mathbb{P}_n^t(I_n)$ is a random distribution.*
We say that the $\{\mu_n\}_{n\in\mathbb{N}}$ converge locally weakly *to $\mathfrak{m}$, a probability measure on $\mathcal{P}(\overline{\mathcal{G}}_*)$, if the law of $\mathbb{P}_n^t(I_n)$ converges weakly to $\mathfrak{m}^t$ for all t.*

Notions of convergence similar to Definition 1.4, and the weaker form of convergence of Definition 2.1 were studied under the name of *metastates for Gibbs measures* (see [3, 24, 34]).
In order to state our main results formally, we need to make few more definitions and state some assumptions:
For any tree T, and for any l positive integer, we denote $\mathsf{T}(l)$ to be the first l generations of T or in other words it is the subtree induced by the vertices of T which are of distances less than equal to l from the root. For each l, consider Ising measures on $\mathsf{T}(l)$ with $(+)$ and $(-)$ boundary conditions:

$$\mu^l_{+,\mathsf{T}}(\underline{x}) \equiv \frac{1}{Z^{l,+}} \exp\Big\{\beta \sum_{(i,j)\in E(\mathsf{T}(l))} x_i x_j + B\sum_{i\in \mathsf{T}(l)} x_i\Big\} \mathbb{I}\big(\underline{x}_{\mathsf{T}\setminus\mathsf{T}(l-1)} = (+)_{\mathsf{T}\setminus\mathsf{T}(l-1)}\big)$$

$$\mu^l_{-,\mathsf{T}}(\underline{x}) \equiv \frac{1}{Z^{l,-}} \exp\Big\{\beta \sum_{(i,j)\in E(\mathsf{T}(l))} x_i x_j + B\sum_{i\in \mathsf{T}(l)} x_i\Big\} \mathbb{I}\big(\underline{x}_{\mathsf{T}\setminus\mathsf{T}(l-1)} = (-)_{\mathsf{T}\setminus\mathsf{T}(l-1)}\big)$$

Here for any $U \subseteq V(\mathsf{T})$, $(+)_U$ (and $(-)_U$) denotes the vector $\{x_i = +1, i \in U\}$ (and $\{x_i=-1, i\in U\}$, respectively). It is well known that as $l\to\infty$ both $\mu^{l,+}$ and $\mu^{l,-}$ converge to probability measures on $\{-1,+1\}^{\mathsf{T}}$, denoted as $\mu^{\beta,B}_{+,\mathsf{T}}$ (plus measure) and $\mu^{\beta,B}_{-,\mathsf{T}}$ (minus measure), respectively (see [26, Chapter IV]), omitting the parameters $\beta,\ B$, when clear from the context. For ν a probability measure, f a function we will sue the shorthand $\langle \nu, f\rangle = \sum_x \nu(x) f(x)$. We will also use $\nu\langle f\rangle$ and some times $\langle f \rangle$ when ν is clear form the context.

For any $\beta, B \geq 0$, and $\mu \in \mathcal{U}$, let

$$U(\beta, B) := \frac{1}{2}\mathbb{E}_\mu\Big[\sum_{i\in\partial\phi} \nu^{\beta,B}_{+,\mathsf{T}}\langle x_\phi \cdot x_i\rangle\Big] = \frac{1}{2}\mathbb{E}_\mu\Big[\sum_{i\in\partial\phi} \nu^{\beta,B}_{-,\mathsf{T}}\langle x_\phi \cdot x_i\rangle\Big] \tag{1.7}$$

Note that the above is well defined and finite if $\mathbb{E}_\mu[\Delta_\phi]$ is finite. Now we are ready to state our first result. This is a generalization of [32, Theorem 2.4.I], for any limiting measure, with a mild continuity assumption.

Theorem 1.5. *Suppose $G_n \stackrel{\text{LWC}}{\Longrightarrow} \mu$ and $U(\,\cdot, 0) \in \mathcal{C}$. Then μ_n converges locally weakly to $\mathfrak{m} = \mu \circ \psi^{-1}$, where $\psi : \mathcal{G}_* \to \mathcal{P}(\overline{\mathcal{G}_*})$ with $\psi(\mathsf{T}) = \delta_{\mathsf{T}} \otimes \frac{1}{2}(\nu_{+,\mathsf{T}} + \nu_{-,\mathsf{T}})$.*

To state the remaining results, we need to define *ergodic unimodular measure*, which we do in the next subsection.

1.3. **Random Walks and Ergodicity.** Given $G \in \mathcal{G}_*$, chosen accordingly $\mu \in \mathcal{U}$, consider the discrete time simple random walk (SRW) starting at the root. By moving the root of the graph G to the current position of the random walk, a random walk on $\mathcal{G}_*$ is induced, which is commonly known as "walk from the point of view of the particle" (see [4, §4]). From [4, Theorem 4.1], for any $\mu \in \mathcal{U}$, when the initial distribution is biased by the degree of the root, the discrete time SRW is reversible and stationary. For every $\mu \in \mathcal{U}$, the probability measure on the trajectory of the Markov chain determined by the environment starting at o with distribution μ, biased by the degree, will be denoted by $\hat{\mu}$.

Clearly the class $\mathcal{U}$ is convex (with $\mu \in \mathcal{U}$ *extremal*, if it cannot be written as a convex combination of other elements in $\mathcal{U}$). From [4, Theorem 4.6, Theorem 4.7] it follows that for extremal $\mu \in \mathcal{U}$, $\hat{\mu}$ is ergodic (i.e. every shift-invariant event is $\hat{\mu}$ trivial). By a slight abuse of notation when $\hat{\mu}$ is ergodic, the corresponding μ will also be termed as ergodic, and will be denoted as μ^e. Expectation of the degree of the root under μ will often be denoted by $\overline{deg}(\mu)$.

To prove the next result we will assume another additional condition on the graph sequence:

Definition 1.6. *A finite graph $G = (V, E)$ is a $(\delta_1, \delta_2, \lambda)$* edge-expander *if, for any set of vertices $S \subseteq V$, with $\delta_1|V| \leq |S| \leq \delta_2|V|$, $|\partial S| \geq \lambda|S|$.*

Now we have the following result, which generalizes [32, Theorem 2.4.II] for ergodic limiting tree measure, and moreover it identifies the set of possible limit points for any general limiting measure.

Theorem 1.7. *Let $G_n \stackrel{\text{LWC}}{\Longrightarrow} \mu$. Assume that for every $0 < \delta < 1/2$, $\{G_n\}_{n\in\mathbb{N}}$ are $(\delta, 1/2, \lambda_\delta)$ edge-expanders for some $\lambda_\delta > 0$, with uniform bounded degrees. Also assume $U(\,\cdot\,, 0) \in \mathcal{C}$ then*
(i) Every subsequential local weak limit of $\mu_{n,+}$ is supported on

$$\{\delta_{\mathsf{T}} \otimes \overline{\nu}_{+,\mathsf{T}}, \mathsf{T} \in \mathcal{G}_*, \overline{\nu}_{+,\mathsf{T}} \in \{\nu_{+,\mathsf{T}}, \frac{1}{2}(\nu_{+,\mathsf{T}} + \nu_{-,\mathsf{T}})\}.$$

If in addition μ is ergodic then
(ii) $\mu_{n,+}$ converges locally weakly to $\mathfrak{m}_+ = \mu \circ \psi_+^{-1}$ where $\psi_+ : \mathcal{G}_* \to \mathcal{P}(\overline{\mathcal{G}}_*)$ with $\psi_+(\mathsf{T}) = \delta_{\mathsf{T}} \otimes \nu_{+,\mathsf{T}}$.

Remark 1.8. To circumvent non-essential technical issues, hereafter $\mu_{n,+}$ denotes the probability measure μ_n conditioned on $\sum_i x_i \geq U$, where $U \sim \text{Unif}(-1/2, 1/2)$ is independent of μ_n. This modification is relevant only for n even, where we now add each $\underline{x}$ with $\sum_i x_i = 0$ with probability $1/2$ to the support of $\mu_{n,+}$ and to that of $\mu_{n,-}$ with probability $1/2$, resulting with $\mu_n = \frac{1}{2}\mu_{n,+} + \frac{1}{2}\mu_{n,-}$ for each n.

Remark 1.9. Recall the example in [32, §2.3], where it is shown that an expander-like condition is necessary to obtain the convergence of $\mu_{n,+}$.

1.4. **Multi-type Galton-Watson** (MGW) **trees.** Now we consider the case when the limiting tree is MGW. Since any measure arising as a limit of local weak convergence is unimodular, we need to define the unimodular version of MGW measure, which will be denoted by UMGW. First we define a *configuration model* below, which will be relevant for UMGW measure:

Definition 1.10. *Consider a finite type space $\mathcal{Q}$. Now for each n define the random graph $G_n = (V_n, E_n)$ as follows: Each $v \in V_n := [n]$, is assigned a type $q(v) \in \mathcal{Q}$ independently according to some distribution θ on $\mathcal{Q}$. Given that a vertex v has type q, there is a kernel $P_q(\cdot) = P(\cdot|q(v))$ which determines a random collection of ordered half-edges, denoted by (v, e_v) and each labeled with some type $q(e_v)$ according to to $P_{q(v)}$. We then obtain the graph by matching half-edges uniformly: for $v, w \in V_n$ a uniform match is permitted only if $q(e_v) = q(w)$ and $q(v) = q(e_w)$.*

Definition 1.11. *For θ any distribution on a finite type space $\mathcal{Q}$, $i, j \in \mathcal{Q}$, and any kernel $P_i(\cdot) := P_i(\underline{k}) = P_i(k_1, k_2, \ldots, k_q)$, let*

$$A(i,j) = \sum_{\underline{k}} P_i(\underline{k}) k_j,$$

where k_l denotes the number of children of type l. Now assume

$$\theta(i)A(i,j) = \theta(j)A(j,i), \text{ for all } i, j \in \mathcal{Q}.$$

Further define kernel

$$\rho_{i,j}(\underline{k}) = P_i(\underline{k} + e_j)\frac{k_j + 1}{A(i,j)},$$

where e_j denotes the vector with 1 at j^{th} co-ordinate and 0 elsewhere. Under the UMGW *measure a tree looks like the following: Type of the root is chosen according to θ. Then conditional on the type of the root, say i_0, it's children are chosen according to $P_{i_0}(\cdot)$, and from next generation onwards, off-spring distribution is chosen according to $\rho_{i,j}$ where i is the type of the current vertex and j being the type of the parent.*

Remark 1.12. The measure UMGW is unimodular, because it can be easily verified that the LWC limit of the configuration model described in Definition 1.10 is the UMGW measure defined above (for a proof in case of $|\mathcal{Q}| = 1$, see [11, Proposition 2.5]). Note for $|\mathcal{Q}| = 1$, the limiting measure is the UGW measure of [4, Example 1.1], and [10, Section 2.1].

From now on, whenever we consider any UMGW measure, we will implicitly assume that it is *positive regular*, and *non-singular* (for definitions see [6, pp. 184]). In Section 5 we prove the following lemma:

Lemma 1.13. *For any* UMGW *measure, one has that $\beta \mapsto U(\beta, 0) \in \mathcal{C}$.*

Thus, upon applying Theorem 1.5 and Theorem 1.7 we immediately obtain the following corollary:

Corollary 1.14. *Let $G_n \stackrel{\text{LWC}}{\Longrightarrow} \mu$ where μ is either an* UGW *or* UMGW *measure. Then*
(a) μ_n converges locally weakly to $\mathfrak{m} = \mu \circ \psi^{-1}$.
(b) If in addition for every $0 < \delta < 1/2$, $\{G_n\}_{n \in \mathbb{N}}$ are $(\delta, 1/2, \lambda_\delta)$ edge-expanders for some $\lambda_\delta > 0$, and have uniform bounded degree then $\mu_{n,+}$ converges locally weakly to $\mathfrak{m}_+ = \mu \circ \psi_+^{-1}$.

As a consequence of our results on UMGW measure, we can easily prove the following interesting lemma, which may be of independent interest:

Lemma 1.15. *Consider any* UGW$(\underline{P})$*. Fix any $\beta_0 > \beta_c$. Let $\{h_i^{0,+}\}$ be i.i.d. copies of $h^{0,+}$ where $h^{0,+}$ is the limit point the recursion (1.8) for $\beta = \beta_0$, started with $+\infty$, and also write $h_1 \succeq h_2$ for*

stochastic domination. Then for any $\beta \ge \beta_0$ consider a sequence random variables $\{h^{(t)}\}$ defined by any random variable $h^{(0)}$ such that $h^{(0)} \succeq h^0$, and for $t > 0$,

$$h^{(t+1)} \stackrel{d}{=} \sum_{i=1}^{K-1} \operatorname{atanh}[\tanh(\beta)\tanh(h_i^{(t)})], \tag{1.8}$$

where $K \sim \underline{\rho}$, and $h_i^{(t)}$ are i.i.d. copies of $h^{(t)}$. Then the sequence of random variables $\{h^{(t)}\}$ converge to the random variable $h^{\beta,+}$ where $h^{\beta,+}$ is the limit of the recursion equation (1.8).

Remark 1.16. Since there is a one-one relation between the fixed point equation, and the *splitting Gibbs measure* (see [12, Remark 1.12, Remark 2.6]), Lemma 1.15 further shows that there is a unique splitting Gibbs measure, with $h \succeq h^{(0)}$ for every $\beta_0 \ge \beta_c$. Similar conclusions can applies for UMGW measures.

Examples of expanders graphs are abundant in literature. Here we are interested in finding a rich class of expander graphs which admit local weak limit, and thus all of our results will be applicable for those graphs. In this direction, it is already well-known that random d-regular graphs are expander graphs. In the lemma below we show that this extends to the configuration model defined in Definition 1.10, and thus for this graph ensemble Corollary 1.14 applies without any extra assumption.

Lemma 1.17. *For every $0 < \delta_0 < 1/2$, there exists a $\lambda_{\delta_0} > 0$ such that the random graph chosen according to the configuration model, defined in Definition 1.10, is an $(\delta_0, 1/2, \lambda_{\delta_0})$ edge-expander with probability tending to 1, whenever the graphs are of uniform bounded degree, and $P_i(|\underline{k}|) = 0$, whenever $|\underline{k}| \in \{0, 1, 2\}$, for every $i \in \mathcal{Q}$.*

Outline of the paper.

- In §2 we show that $\{\mu_n\}$ and $\{\mu_{n,+}\}$ have subsequential limits. As a first step we establish an weaker notion of convergence, termed as *convergence on average*, and then extend it to the local weak convergence. It is further shown that the convergence on average limit is an Ising Gibbs measure, and the local weak limit is supported on Ising Gibbs measures.
- In §3 using local weak convergence of the graphs we first find the limiting values of a functional of the spin configuration. Then extending [32, Lemma 3.2] we conclude that the limits obtained in §2 are convex combination of plus and minus measure, (for local weak limit the measure is supported on the line joining plus and minus measure). Now using symmetry of μ_n in the sign of $\underline{x}$ we obtain Theorem 1.5.
- In §4 we prove Theorem 1.7. As a first step some functionals are chosen, and shown that by local weak convergence of the graphs, and that of $\mu_{n,+}$ value of the functional converge to that of the limiting trees in expectation. Then using properties of SRW on trees, all but two choices are eliminated. This gives the result for any general limiting tree measure. When the limiting tree measure is ergodic, the same proof yields the desired result.
- In §5 we prove that $\beta \mapsto U(\beta, 0) \in \mathcal{C}$. Two cases $\beta \in (\beta_c, \infty)$ and $\beta = \beta_c$ are done separately in Lemma 5.3 and Lemma 5.6 respectively. To prove the continuity of $U(\beta, 0)$in $\beta > \beta_c$ we come up with some a sequence of random variables that increases up to the root magnetization under $\nu_{+,\mathsf{T}}$. $\beta = \beta_c$ case is done by using a capacity criterion. We further confirm that the assumptions needed for both the lemmas are satisfied for UMGW measures, and thus we obtain Corollary 1.14. Lastly we prove Lemma 1.17.

Acknowledgements. We thank Allan Sly for proposing the choice of functionals in (4.1), Andrea Montanari for a key idea in the proof of Lemma 5.3, and Yuval Peres for helpful discussions on

the proof of Lemma 5.6. We also thank Noga Alon, Russell Lyons, and Nike Sun for many helpful conversations.

2. Convergence to Ising Gibbs measure

In this section we show that subsequential local weak limits exist for the sequences $\{\mu_n\}$ and $\{\mu_{n,+}\}$. As first step we prove such results for *convergence on the average*, a weaker form of convergence, which we define below:

Definition 2.1. *Denoting* $\mathbb{P}_n^t = \mathbb{E}_{U_n}(\mathbb{P}_n^t(I_n))$ *to be the average of* $\mathbb{P}_n^t(I_n)$*, we say* $\{\mu_n\}_{n\in\mathbb{N}}$ *converge* on average *to* ν*, a probability measure on* $(\overline{\mathcal{G}}_*, \mathscr{C}_{\overline{\mathcal{G}}_*})$*, if for every* $t > 0$*,*

$$\mathbb{P}_n^t \Rightarrow \nu^t, \ \text{ as } n \to \infty. \tag{2.1}$$

Note that if μ_n converge locally weakly then it automatically implies convergence on average.

Remark 2.2. A stronger notion of convergence appeared in [32, Definition 2.3], where μ_n are said to converge *in probability* to $\nu \in \mathcal{P}(\overline{\mathcal{G}}_*)$, if for every $t > 0$,

$$\lim_{n\to\infty} U_n(d_{TV}(\mathbb{P}_n^t(I_n), \overline{\nu}^t) > \varepsilon) = 0. \tag{2.2}$$

This is equivalent to local weak convergence when $\mathfrak{m} = \delta_\nu$. Unlike [32] it does not appear here due to the randomness in the limiting tree.

Lemma 2.3. *(i) Suppose* $G_n \stackrel{\text{LWC}}{\Longrightarrow} \mu$*. Then for any subsequence* $\{n_m\}_{m\in\mathbb{N}}$ *there exists a further subsequence* $\{n_{m_k}\}_{m\in\mathbb{N}}$ *such that* $\mu_{n_{m_k}}$ *converges locally weakly on average to* $\overline{\nu} \in \mathcal{P}(\overline{\mathcal{G}}_*)$*, where* $\overline{\nu}$ *might depend on the choice of the subsequence.*

(ii) Same conclusion holds for the sequence of probability measures $\{\mu_{n,+}\}_{n\in\mathbb{N}}$*, denoting tis subsequential limit by* $\overline{\nu}_+$*.*

Proof: Note that for every t, $\{\mathbb{P}_n^t\}_{n\in\mathbb{N}}$ are probability measures on $\overline{\mathcal{G}}_*(t)$, which is a discrete space. Further note that, For every $G \in \mathcal{G}_*(t)$,

$$\mathbb{P}_n^t(G, \{-1,1\}^G) = B_{I_n}(t)(G) := \sum_{i=1}^n \mathbb{I}(B_i(t) \simeq G).$$

The convergence of $\{U(G_n)\}_{n\in\mathbb{N}}$ implies that of $\{B_{I_n}(t)\}$, hence $\{B_{I_n}(t)\}$ are tight. Any compact subset of $\mathcal{G}_*(t)$ is finite, and per graph $G \in \mathcal{G}_*(t)$, space of all marks $\{-1,1\}^G$, is finite. Thus we obtain a finite subset of $\overline{\mathcal{G}}_*(t)$, outside which, under $\mathbb{P}_n^t$ probability is small, for all n. Therefore $\{\mathbb{P}_n^t\}$ are uniformly tight and hence relatively compact. So we get a subsequential limit, say $\overline{\nu}_t$. Now proof will be complete by taking a diagonal subsequence and using Kolmogorov's extension theorem, once we show that

$$\overline{\nu}_t(G_0, \underline{x}_{G_0}) = \sum_{\substack{G\in\mathcal{G}_*(t+1)\\ G(t)\simeq G_0}} \sum_{\underline{x}_{G(t,t+1)}} \overline{\nu}_{t+1}(G, \underline{x}_G) \tag{2.3}$$

holds for every $G_0 \in \mathcal{G}_*(t)$, $\underline{x}_{G_0} \in \{-1,1\}^{G_0}$ and every fixed $t > 0$. Noting that the expression (2.3) holds when $\overline{\nu}_t$ and $\overline{\nu}_{t+1}$ are replaced by $\mathbb{P}_n^t$ and $\mathbb{P}_n^{t+1}$ respectively, completes the proof for $\{\mu_n\}_{n\in\mathbb{N}}$. Proof for $\{\mu_{n,+}\}_{n\in\mathbb{N}}$ is similar, details are omitted. □

Next we extend the above lemma strengthening the convergence to average to local weak convergence.

Lemma 2.4. *(i) Let* $G_n \stackrel{\text{LWC}}{\Longrightarrow} \mu$. *Then for any subsequence* $\{n_m\}_{m\in\mathbb{N}}$ *there exists a further subsequence* $\{n_{m_k}\}_{k\in\mathbb{N}}$ *such that* $\mu_{n_{m_k}}$ *converges locally weakly to* $\mathfrak{m}$*, a probability measure on* $\mathcal{P}(\overline{\mathcal{G}}_*)$.

(ii) Same conclusion holds for $\{\mu_{n,+}\}_{n\in\mathbb{N}}$*, denoting its subsequential limit by* $\mathfrak{m}_+$.

Proof: All that we need to show is the uniform tightness of $\{\mathbb{P}_n^t(I_n)\}$ for every positive integer t. Then rest of the proof follows by taking a diagonal subsequence and noting that the subsequential local weak limits obtained for different t's are consistent.

To this end note that from the proof of Lemma 2.3 we obtain a finite set $\mathcal{G}_\varepsilon \subset \mathcal{G}_*(t)$, such that $\inf_n B_{I_n}(t)(\mathcal{G}_\varepsilon) \ge 1-\varepsilon$. As $\mathcal{G}_\varepsilon$ is finite, so is the set $\{(G, \underline{x}_G), G \in \mathcal{G}_\varepsilon, \underline{x}_G \in \{-1,1\}^G\}$ and thus by Prohorov's theorem the set of all measures on it is compact. Hence the set $\mathcal{M}_\varepsilon := \{\delta_G \otimes \mu_G, G \in \mathcal{G}_\varepsilon, \mu_G \in \mathcal{P}(\{-1,1\}^G)\}$ is pre-compact under the usual topology. Now from $\inf_n B_{I_n}(t)(\mathcal{G}_\varepsilon) \ge 1-\varepsilon$. $\mathcal{G}_\varepsilon$ it readily follows $\inf_n \mathbb{P}_n^t(I_n)(\mathcal{M}_\varepsilon) \ge 1-\varepsilon$, proving the uniform tightness of $\{P_n^t(I_n)\}$. Proof for $\{\mu_{n,+}\}$ is similar, details omitted. □

If a sequence of Gibbs measures converge weakly to some probability measure then it is well-known that the limiting measure is also a Gibbs measure. Here $\{\mu_n\}_{n\in\mathbb{N}}$ are Gibbs measures and thus it is natural to expect that the subsequential limits will also be Gibbs measures. Since the limiting graph is non-deterministic, we need to uplift the definition of Gibbs measure. Before defining Gibbs measure we need the following definitions:

Note that there is a one-to-one correspondence between $G \in \mathcal{G}_*$ and $\{(G, x_G),\ x_G \in \{-1,1\}^G\} \subset \overline{\mathcal{G}}_*$. Thus there is a natural embedding of $\mathscr{C}_{\mathcal{G}_*}$ in $\mathscr{C}_{\overline{\mathcal{G}}_*}$, where $\mathscr{C}_{\mathcal{G}_*}$ is identified with the σ-algebra generated by $\{B_G(r)\}_{r\in\mathbb{N},\ G\in\mathcal{G}_*}$, and $B_G(r) = \cup_{\underline{x}_G} B_{(G,\underline{x}_G)}(r)$. With this embedding in mind, for any $\nu \in \mathcal{P}(\overline{\mathcal{G}}_*)$, we define,

$$\nu_G(\cdot) := \nu\big((G,\cdot)\big|G\big) = \nu\big((G,\cdot)\big|\mathscr{C}_{\mathcal{G}_*}\big). \tag{2.4}$$

Since $\overline{\mathcal{G}}_*$ is a polish space the conditional probability above is well defined and can be taken to be a version of the regular conditional probability measure (see [38, §9.2]). Now we are ready to define notion of Gibbs measure, for $\nu \in \mathcal{P}(\overline{\mathcal{G}}_*)$.

Definition 2.5. *A probability measure* $\nu \in \mathcal{P}(\overline{\mathcal{G}}_*)$ *will be called on Ising Gibbs measure, if for* ν*-a.e* G*,* $\nu_G(\ \cdot\ )$ *is an Ising Gibbs measure.*

Remark 2.6. A measure on an infinite graph, is said to be an Ising Gibbs measure if it satisfies DLR condition (see [18, Chapter 2]). Thus $\nu \in \mathcal{P}(\overline{\mathcal{G}}_*)$ will be an Ising Gibbs measure if, ν- a.e. G, if we have,

$$\nu_G(\underline{x}_{G(t)}|\underline{x}_{G(t,\infty)}) = \frac{\exp\left\{\beta \sum_{(i,j)\in E(G(t+1))} x_i x_j\right\}}{\sum_{\underline{x}_{G(t)}} \exp\left\{\beta \sum_{(i,j)\in E(G(t+1))} x_i x_j\right\}}, \tag{2.5}$$

where $G(t,\infty) = G\backslash G(t)$. To make sense of the conditional probability above, we need to define an appropriate sub σ-algebra, say $\mathscr{C}_{\mathcal{G}_*^t}$. Note that this sub σ-algebra can be generated by $\{B_G^t(r)\}$, where $B_G^t(r) = \cup_{\underline{x}_{G(\min\{r,t\})}} B_{(G,\underline{x}_G)}(r)$, and $\nu_G(\ \cdot\ |\underline{x}_{G(t,\infty)})$ will be a version of regular conditional probability measure of ν given the sub σ-algebra $\mathscr{C}_{\mathcal{G}_*^t}$. Further note that, upon defining appropriate

sub σ-algebras, for ν to be an Ising Gibbs measure, it is enough to show,

$$\nu(\underline{x}_{G(t)}|G(\overline{t}),\ \underline{x}_{G(t,t_+)}) = \frac{\exp\left\{\beta \sum_{(i,j)\in E(G(t+1))} x_i x_j\right\}}{\sum_{\underline{x}_{G(t)}} \exp\left\{\beta \sum_{(i,j)\in E(G(t+1))} x_i x_j\right\}}. \tag{2.6}$$

for all $\overline{t} \geq t_+ > t$, where $G(t,t_+) = G(t_+)\backslash G(t)$.

Next we identify all subsequential limits of μ_n and $\mu_{n,+}$ as Ising Gibbs measure.

Lemma 2.7. *(i) Every subsequential limit $\overline{\nu}$, of μ_n must be an Ising Gibbs measure. Moreover every subsequential local weak limit $\mathfrak{m}$ is supported on the space of all Ising Gibbs measures.*

(ii) If the degrees of $\{G_n\}$ are uniformly bounded then same conclusions hold for $\mu_{n,+}$.

Proof: When ν is supported on trees, note establishing (2.6) for $\overline{t} = t_+ = t+1$ is sufficient. Further noting that the space required to define $\nu(\underline{x}_{G(t)}|G(\overline{t}),\ \underline{x}_{G(t,t_+)})$ is discrete, by the definition of subsequential limit,

$$\overline{\nu}(\underline{x}_{G(t)}|G(t+1),\ \underline{x}_{G(t,t+1)}) = \frac{\overline{\nu}\{(G(t+1),\underline{x}_{G(t+1)})\}}{\overline{\nu}\{(G(t+1),\underline{x}_{G(t,t+1)})\}} = \lim_{m\to\infty} \frac{\mathbb{P}^{t+1}_{n_m}\{(G(t+1),\underline{x}_{G(t+1)})\}}{\sum_{\underline{x}_{G(t)}} \mathbb{P}^{t+1}_{n_m}\{(G(t+1),\underline{x}_{G(t+1)})\}}, \tag{2.7}$$

For ease of writing we introduce a few more notation: For any vertex $i \in G_n$, denote $B_i^c(t) = V_n\backslash B_i(t)$ and $E_i(t)$ will denote the edge set in the subgraph $B_i(t)$ and $E_i^c(t) = E_n\backslash E_i(t)$. Now further denote:

$$\begin{aligned}
&m_{i,t+1}(\underline{x}) = \sum_{j\in B_i(t+1)} x_j,\ m_{i^c,t+1}(\underline{x}) = \sum_{j\in B_i^c(t+1)} x_j,\\
&\psi_{E_i(t+1)}(\underline{x}) = \exp\Big[\sum_{(k,l)\in E_i(t+1)} x_k x_l\Big],\ \psi_{E_i^c(t+1)}(\underline{x}) = \exp\Big[\sum_{(k,l)\in E_i^c(t+1)} x_k x_l\Big],\\
&\psi_{G(t+1)}(\underline{x}_{G(t+1)}) = \exp\Big[\sum_{(k,l)\in E(G(t+1))} x_k x_l\Big],\ Z_i^0(\underline{x}_{\partial B_i(t+1)}) = \sum_{\underline{x}_{B_i^c(t+1)}} \psi_{E_i^c(t+1)}(\underline{x}).
\end{aligned}$$

When $B_i(t+1) \simeq G(t+1)$, $\psi_{E_i(t+1)}(\underline{x})$ is free of n and i, and depends only on $\underline{x}_{G(t+1)}$, thus

$$\mathbb{P}^{t+1}_n\{(G(t+1),\underline{x}_{G(t+1)})\} = \frac{\psi_{G(t+1)}(\underline{x}_{G(t+1)})}{Z_n} \times \frac{1}{n}\Big[\sum_{i=1}^n \mathbb{I}\Big(B_i(t+1)\simeq G(t+1)\Big) Z_i^0(\underline{x}_{\partial B_i(t+1)})\Big] \tag{2.8}$$

and similarly for the denominator,

$$\sum_{\underline{x}_{G(t)}} \mathbb{P}^{t+1}_n\{(G(t{+}1),\underline{x}_{G(t+1)})\} = \sum_{\underline{x}_{G(t)}} \frac{\psi_{G(t+1)}(\underline{x}_{G(t+1)})}{Z_n} \times \frac{1}{n}\Big[\sum_{i=1}^n \mathbb{I}\Big(B_i(t{+}1)\simeq G(t{+}1)\Big) Z_i^0(\underline{x}_{\partial B_i(t+1)})\Big]. \tag{2.9}$$

Dividing (2.8) by (2.9), we note that $\mathbb{P}^t_n$'s are Ising Gibbs measure and hence so is the subsequential limit.

Now to prove $\mathfrak{m}$ is supported on Ising Gibbs measure, consider the set of all measures ν such that (2.6) is satisfied for a fixed $\overline{t} \geq t_+ > t$, call this set $\mathcal{M}$. It is enough to show that $\mathfrak{m}$ is supported on this set, as the proof then can be competed by taking intersection over $\overline{t} \geq t_+ > t$.

Note $\mathbb{P}_n^{\overline{t}}(I_n)$ is supported on $\{\delta_{B_i(\overline{t})} \otimes \mu_{n,B_i(\overline{t})}\}$ and whenever $B_i(\overline{t}) \simeq G$, for some $G \in \mathcal{G}_\varepsilon$ ($\mathcal{G}_\varepsilon$ defined in the proof of Lemma 2.4), $\mu_n(\underline{x}_{B_i(t)}|\underline{x}_{B_i(t,t_+)})$ free of n, i and satisfies (2.6). Thus on $\mathcal{M}_\varepsilon$ (defined in Lemma 2.4), $\mathbb{P}_n^{\overline{t}}(I_n)$ is supported on those measures only, which belong to $\mathcal{M}$. Now further noting that $\mathcal{M}$ is closed,

$$1-\varepsilon \leq \limsup_{m\to\infty} \mathbb{P}_{n_m}^{\overline{t}}(I_{n_m})(\mathcal{M}_\varepsilon) = \limsup_{m\to\infty} \mathbb{P}_{n_m}^{\overline{t}}(I_{n_m})(\mathcal{M}) \leq \mathfrak{m}^{\overline{t}}(\mathcal{M}).$$

Since the above holds for every $\varepsilon > 0$, we have $\mathfrak{m}^{\overline{t}}(\mathcal{M}) = 1$. Now whether ν satisfies (2.6) or not, depends only the marginal ν^t we conclude $\mathfrak{m}(\mathcal{M}) = 1$.

Now to prove the result, for $\mu_{n,+}$ we note that at finite n, it is not a Gibbs measure because of conditioning on $\sum_i x_i > 0$. However we will argue that the effect of conditioning washes away in the limit.

To this end note that the expressions (2.8) and (2.9) remains valid for$\mu_{n,+}$ when $Z_i^0(\underline{x}_{\partial B_i(t+1)})$ is replaced by $Z_i(\underline{x}_{B_i(t+1)})$, where

$$Z_i(\underline{x}_{B_i(t+1)}) = \sum_{\underline{x}_{B_i^c(t+1)}} \psi_{E_i^c(t+1)}(\underline{x}) \mathbb{I}\Big(m_{i^c,t+1}(\underline{x}) > -m_{i,t+1}(\underline{x})\Big).$$

As

$$\mathbb{I}\Big(m_{i^c,t+1}(\underline{x}) > |B_i(t+1)|\Big) \leq \mathbb{I}\Big(m_{i^c,t+1}(\underline{x}) > -m_{i,t+1}(\underline{x})\Big) \leq \mathbb{I}\Big(m_{i^c,t+1}(\underline{x}) > -|B_i(t+1)|\Big),$$

and letting

$$\mu^*(\underline{x}_{G(t)}|\underline{x}_{G(t,t+1)}) := \frac{\psi_{G(t+1)}(\underline{x}_{G(t+1)})\frac{1}{n}\sum_{i=1}^n \mathbb{I}\Big(B_i(t+1)\simeq G(t+1)\Big)}{\sum_{\underline{x}_{G(t)}} \psi_{G(t+1)}(\underline{x}_{G(t+1)})\frac{1}{n}\sum_{i=1}^n \mathbb{I}\Big(B_i(t+1)\simeq G(t+1)\Big)} = \frac{\psi_{G(t+1)}(\underline{x}_{G(t+1)})}{\sum_{\underline{x}_{G(t)}} \psi_{G(t+1)}(\underline{x}_{G(t+1)})},$$

we note that

$$\mu^*(\underline{x}_{G(t)}|\underline{x}_{G(t,t+1)}) \times L_n \leq \frac{\mathbb{P}_{n,+}^{t+1}\{(G(t+1), \underline{x}_{G(t+1)})\}}{\sum_{\underline{x}_{G(t)}} \mathbb{P}_{n,+}^{t+1}\{(G(t+1), \underline{x}_{G(t+1)})\}} \leq \mu^*(\underline{x}_{G(t)}|\underline{x}_{G(t,t+1)}) \times L_n^{-1}, \tag{2.10}$$

where

$$L_n = \min_{\underline{x}_{\partial G(t+1)}, \{i\in[n]\,:\, B_i(t+1)\simeq G(t+1)\}} \frac{Z_i^-(\underline{x}_{\partial B_i(t+1)})}{Z_i^+(\underline{x}_{\partial B_i(t+1)})},$$

$$\text{and} \quad Z_i^\pm(\underline{x}_{\partial B_i(t+1)}) = \psi_{E_i^c(t+1)}(\underline{x})\mathbb{I}\Big(m_{i^c,t+1}(\underline{x}) > \mp|B_i(t+1)|\Big).$$

Now proof will be done if we can show that $L_n \to 1$ as $\mu^*(\underline{x}_{G(t)}|\underline{x}_{G(t,t+1)})$ satisfies (2.6). To this end define,

$$\hat{\mu}_n^i(\underline{x}_{B_i^c(t+1)}) = \frac{\psi_{E_i^c(t+1)}}{\sum_{\underline{x}_{B_i^c(t+1)}} \psi_{E_i^c(t+1)}}$$

Thus

$$1-\frac{Z_i^-(\underline{x}_{\partial B_i(t+1)})}{Z_i^+(\underline{x}_{\partial B_i(t+1)})}=\frac{\hat{\mu}_n^i\Big(m_{i^c,t+1}(\underline{x})>-|B_i(t+1)|\Big)-\hat{\mu}_n^i\Big(m_{i^c,t+1}(\underline{x})>|B_i(t+1)|\Big)}{\hat{\mu}_n^i\Big(m_{i^c,t+1}(\underline{x})>-|B_i(t+1)|\Big)}. \tag{2.11}$$

Since degrees are uniformly bounded by some number Δ, we have that,

$$\hat{\mu}_n^i(\underline{x}_{B_i^c(t+1)})\geq \exp(-2\beta\Delta|B_i(t+1)|)\mu_n(\underline{x}_{B_i^c(t+1)}).$$

Now using symmetry of the measure $\mu_n(\,\cdot\,)$ with respect to the sign of $\underline{x}$,

$$\hat{\mu}_n^i\Big(m_{i^c,t+1}(\underline{x})>-|B_i(t+1)|\Big)\geq\frac{1}{2}\exp(-2\beta\Delta|B_i((t+1)|).$$

Thus the denominator of (2.11) is bounded away from zero. To control the numerator we note that,

$$\begin{aligned}\hat{\mu}_n^i\Big(m_{i^c,t+1}(\underline{x})>-|B_i(t+1)|\Big)-\hat{\mu}_n^i\Big(m_{i^c,t+1}(\underline{x})>|B_i(t+1)|\Big) &\leq \hat{\mu}_n^i\Big(|m_{i^c,t+1}(\underline{x})|\leq|B_i(t+1)|\Big)\\ &\leq \frac{2C|B_i(t+1)|}{\sqrt{n-|B_i(t+1)|}}\to 0,\end{aligned}$$

where the last inequality follows by an use of [32, Lemma 4.1] and noting that for a graph whose degrees are uniformly bounded by a constant, size of the maximal independent set is proportional to number of vertices.

Now it remains to argue that any subsequential limit $\mathfrak{m}_+$ of $\mu_{n,+}$ is supported on Ising Gibbs measure. Proof can be done similarly as in the case of $\mathfrak{m}$. Changes are needed, as $\mu_{n,+}(\underline{x}_{B_i(t)}|\underline{x}_{B_i(t,t_+)})$ no longer satisfy (2.6), even when $B_i(\overline{t})$ have no cycles. However from (2.10) it is evident that, for any fixed $\delta>0$, for large n, $\mu_{n,+}(\underline{x}_{B_i(t)}|\underline{x}_{B_i(t,t_+)})$ satisfies (2.6) with an error δ, i.e. the ratio of LHS and RHS of (2.6) are bounded below and above by $(1-\delta)$ and $(1+\delta)$. Let $\mathcal{M}^\delta$ denote the set of measures such that (2.6) hold with an error δ, then proceeding as in the case of μ_n, we obtain for every $\delta>0$, $\mathfrak{m}_+(\mathcal{M}^\delta)=1$. Now taking intersection over δ, $t<t_+\leq\overline{t}$ required result is obtained. □

3. Identifying the limit Gibbs measure

Lemma 3.1. *Let $\{G_n\}_{n\in\mathbb{N}}$ be a sequence of graphs such that $G_n\overset{\text{LWC}}{\Longrightarrow}\mu$. Then under the assumption $U(\,\cdot,0)\in\mathcal{C}$,*

$$\lim_{n\to\infty}\frac{1}{n}\sum_{(i,j)\in E_n}\mu_{n,+}\langle x_i\cdot x_j\rangle=\lim_{n\to\infty}\frac{1}{n}\sum_{(i,j)\in E_n}\mu_n\langle x_i\cdot x_j\rangle \;=\; U(\beta,0) \tag{3.1}$$

We use in our proof [15, Lemma 3.1] (a more restricted version of which was key to [10]). We note that [15, Lemma 3.1] is trivial for finite trees. Hereafter we adopt the notation $\mathsf{T}_{x\to y}$ for connected component of the subtree of T rooted at x, after the path between x and y is deleted. During our proof it will be useful to consider vertex dependent magnetic fields B_i, that is, to replace the model (1.1) by

$$\mu(\underline{x})=\frac{1}{Z(\beta,\underline{B})}\exp\Big\{\beta\sum_{(i,j)\in E}x_ix_j+\sum_{i\in V}B_ix_i\Big\}. \tag{3.2}$$

The first result on ferromagnetic Ising models that are heavily used is the Griffiths inequality (see [26, Theorem IV.1.21]).

Proposition 3.2. [Griffith's inequality] *Consider two Ising models $\mu(\cdot)$ and $\mu'(\cdot)$ on graphs $G = (V, E)$ and $G' = (V, E')$, inverse temperatures β and β', and magnetic fields $\{B_i\}$ and $\{B_i'\}$, respectively. If $E \subseteq E'$,, $\beta \le \beta'$ and $0 \le B_i \le B_i'$, for all $i \in V$, then $0 \le \langle \mu, \prod_{i\in U} x_i\rangle \le \langle \mu', \prod_{i\in U} x_i\rangle$ for any $U \subseteq V$, if $|V| < \infty$.*

During our proofs we will sometimes need to work with the marginal of an Ising measure on a tree. Below is an excellent result from [10], which is also used in our proofs:

Proposition 3.3. [10, Lemma 4.1] *For a subtree U of a finite tree T, let $\partial_* U$ denote the subset of vertices U connected by an edge to $W \equiv T\backslash U$ and for each $u \in \partial_* U$ let $\langle x_u\rangle_W$ denote the root magnetization of the Ising model on the maximal subtree T_u of $W \cup \{u\}$ rooted at u. The marginal on U of the Ising measures on T, denoted μ_U^T is then an Ising measure on U with magnetic field $B_u' = \operatorname{atanh}(\langle x_u\rangle_W) \ge B_u$ for $u \in \partial_* U$ and $B_u' = B_u$ for $u \notin \partial_* U$.*

Proof: Clearly for any $t \ge 0$, $\nu_{+,\mathsf{T}}^{\beta,B,t}\langle x_\phi \cdot x_i\rangle$ is continuous in both β, B. Now using Griffith's inequality, $\nu_{+,\mathsf{T}}^{\beta,B,t}\langle x_\phi \cdot x_i\rangle$ is decreasing in t and increasing in β, B for $\beta, B \ge 0$. Thus, using DCT and interchanging the limit in t and β, B, the function $U(\beta, B)$ is right continuous for $\beta, B \ge 0$. From [12], $\phi_n(\beta, B) \to \phi(\beta, B)$ as $n \to \infty$ for $\beta \ge 0$ and $B \in \mathbb{R}$. Since $\beta \mapsto \phi_n(\beta, B)$ are convex functions for each B, so is the limiting function $\phi(\beta, B)$. Thus $\frac{\partial}{\partial\beta}\phi_n(\beta, B) \to \frac{\partial}{\partial\beta}\phi(\beta, B)$ on a co-countable set. Hereafter convergence of f_n to f on co-countable set will be denoted by $f_n(\ \cdot\ ) \overset{\mathbb{Q}^c}{\to} f(\ \cdot\ )$ and $f(\ \cdot\ ) \overset{\mathbb{Q}^c}{=} g(\ \cdot\ )$ when f and g agree on a co-countable set. We further claim that for every $B > 0$,

$$\lim_{n\to\infty} \frac{\partial}{\partial\beta}\phi_n(\beta, B) = \lim_{n\to\infty} \mathbb{E}_n\Big[\sum_{(i,j)\in E_n} \mu_n\langle x_i \cdot x_j\rangle\Big] = U(\beta, B). \tag{3.3}$$

To see this first note that by Griffith's inequality and local weak convergence, for every $t \ge 2$,

$$\mathbb{E}_\mu\Big[\frac{1}{2}\sum_{i\in\partial\phi} \nu_{0,\mathsf{T}}^t\langle x_\phi \cdot x_i\rangle\Big] \le \liminf_{n\to\infty} \frac{\partial}{\partial\beta}\phi_n(\beta, B) \le \limsup_{n\to\infty} \frac{\partial}{\partial\beta}\phi_n(\beta, B) \le \mathbb{E}_\mu\Big[\frac{1}{2}\sum_{i\in\partial\phi} \nu_{+,\mathsf{T}}^t\langle x_\phi \cdot x_i\rangle\Big]$$

By Proposition 3.3, using similar ideas as in [10], for any $i \in \partial\phi$, $\nu_{+/0,\mathsf{T}}^t\langle x_\phi \cdot x_i\rangle = \overline{\mu}\langle x_1 \cdot x_2\rangle$, where

$$\overline{\mu}(x_1, x_2) = \frac{1}{z_{12}} \exp\{\beta x_1 x_2 + H_1 x_1 + H_2 x_2\}$$

and $H_1 = \operatorname{atanh}[m_{+/f}^t(\mathsf{T}_{\phi\to i})]$, $H_2 = \operatorname{atanh}[m_{+/f}^{t-1}(\mathsf{T}_{i\to\phi})]$. Since $\overline{\mu}\langle x_1 \cdot x_2\rangle$ is a continuous function of H_1 and H_2, upon applying [15, Lemma 3.1] for $m_{+/f}^t(\mathsf{T}_{\phi\to i})$ and $m_{+/f}^{t-1}(\mathsf{T}_{i\to\phi})$ and recalling that $\mathbb{E}_\mu[\Delta_\phi] < \infty$, we get by DCT that,

$$\lim_{t\to\infty} \mathbb{E}_\mu\Big[\frac{1}{2}\sum_{i\in\partial\phi} \nu_{0,\mathsf{T}}^t\langle x_\phi \cdot x_i\rangle\Big] = \lim_{t\to\infty} \mathbb{E}_\mu\Big[\frac{1}{2}\sum_{i\in\partial\phi} \nu_{+,\mathsf{T}}^t\langle x_\phi \cdot x_i\rangle\Big].$$

This establishes (3.3) and consequently for every $B > 0$, $\frac{\partial}{\partial\beta}\phi(\beta, B) \overset{\mathbb{Q}^c}{=} U(\beta, B)$. Now fixing a sequence $\{B_m\}$ such that $B_m \downarrow 0$, using convexity property of $\beta \mapsto \phi(\beta, B)$, we have $\frac{\partial}{\partial\beta}\phi(\beta, B_m) \overset{\mathbb{Q}^c}{\to} \frac{\partial}{\partial\beta}\phi(\beta, 0)$. Since $B \mapsto U(\beta, B)$ is right continuous, $U(\beta, 0) \overset{\mathbb{Q}^c}{=} \frac{\partial}{\partial\beta}\phi(\beta, 0)$.

Now noting $\frac{\partial}{\partial\beta}\phi_n(\beta, 0) \overset{\mathbb{Q}^c}{\to} \frac{\partial}{\partial\beta}\phi(\beta, 0)$, we have $\frac{\partial}{\partial\beta}\phi_n(\beta, 0) \overset{\mathbb{Q}^c}{\to} U(\beta, 0)$. By assumption $U(\ \cdot, 0) \in \mathcal{C}$ is continuous at β, hence said convergence hold for *all* β. First equality in (3.1) in follows from the facts that $\mu_n = \frac{1}{2}\mu_{n,+} + \frac{1}{2}\mu_{n,-}$ and for each $\underline{x}$, $\mu_{n,+}(\underline{x}) = \mu_{n,-}(-\underline{x})$. ☐

Lemma 3.4. *For any $\mu \in \mathcal{U}$ and any collection of Ising Gibbs measures ν_{T} on T,*

$$\mathbb{E}_\mu\Big[\sum_{i\in\partial\phi}\nu_{\mathsf{T}}\langle x_\phi\cdot x_i\rangle\Big] \le \mathbb{E}_\mu\Big[\sum_{i\in\partial\phi}\nu_{+,\mathsf{T}}\langle x_\phi\cdot x_i\rangle\Big] = \mathbb{E}_\mu\Big[\sum_{i\in\partial\phi}\nu_{-,\mathsf{T}}\langle x_\phi\cdot x_i\rangle\Big], \tag{3.4}$$

and the inequality is strict unless μ a.e. every T, ν_{T} is a convex combination of $\nu_{+,\mathsf{T}}$ and $\nu_{-,\mathsf{T}}$.

The proof below, and some of the proofs later use a notion, termed *branching number*, defined for infinite trees. There are close connections between branching number, and recurrence/ transience properties of SRW, and also phase transitions of Ising models on trees (e.g. see [27, 28]), which we have used in our proofs. For completeness we define branching number of a tree below: Heuristically branching number of tree is the average number of branches coming out from each node. Formally, denoting branching number of tree T by $\mathrm{br}\,\mathsf{T}$,

$$\mathrm{br}\,\mathsf{T} = \Big\{\lambda > 0 : \inf_{\Pi}\sum_{\sigma\in\Pi}\lambda^{-|\sigma|} = 0\Big\},$$

where Π is a cutset i.e. a finite set of vertices such that every infinite path from root would intersect it and $|\sigma|$ is the distance of the shortest path to σ from root.
Proof: By Griffiths inequality for any tree T,

$$\sum_{i\in\partial\phi}\nu_{\mathsf{T}}\langle x_\phi\cdot x_i\rangle \le \sum_{i\in\partial\phi}\nu_{+,\mathsf{T}}\langle x_\phi\cdot x_i\rangle. \tag{3.5}$$

If T is finite then there is only one Gibbs measure and hence equality holds. Denote A_0 to be set of all finite trees. Thus if $\mu(A_0) = 1$ then the above is trivially true. So assume $\mu(A_0) < 1$. Noting that A_0 is invariant under non-rooted isomorphism, $\mu_0 := \mu(\,\cdot\,|A_0^c)$ is also a unimodular measure, supported on infinite trees. In the rest of the proof we will work with μ_0.
Now noting that Given a T, since the space of all Ising Gibbs measure is convex, any Ising Gibbs measure is a mixture of extremal Ising Gibbs measure. Let A_1 be the space of all trees T such that for any extremal measure $\nu_{\mathsf{T}} \notin \{\nu_{+,\mathsf{T}}, \nu_{-,\mathsf{T}}\}$ strict inequality in (3.5) holds. Let A_2 be the set of trees for which there exist at least one extremal measure other than $\nu+, \mathsf{T}$ and $\nu_{-,\mathsf{T}}$ for which in (3.5) equality holds. Since any Ising Gibbs measure on a T is a mixture of extremal Ising Gibbs measure on T (see [18, Chapter 7]), proof will be completed once we establish $\mu_0(A_1) > 0$. Suppose not, i.e. $\mu_0(A_2) = 1$.
Now consider any infinite tree T. Fix any extremal Ising Gibbs measure ν_{T} on T. Now for any $i, j \in V(\mathsf{T})$, such that $i \sim j$, define

$$m^{\nu}_{i\to j} := \lim_{l\to\infty}\mathbb{E}_{\mathsf{T}_{i\to j}}(x_i|\underline{x}_{\mathsf{T}_{i\to j}(l,\infty)}) = \nu_{\mathsf{T}_{i\to j}}\langle x_i\rangle,$$

where $\mathbb{E}_{\mathsf{T}_{i\to j}}$ denotes the expectation with respect to the Ising measure on $\mathsf{T}_{i\to j}$ and the boundary condition $\underline{x}_{\mathsf{T}_{i\to j}(l,\infty)}$ according to ν_{T}. Existence of the limit is guaranteed by the backward martingale convergence theorem and by tail triviality of extremal measure ν_{T} (see [18, Chapter 7]) the limits are a.s. constant and equal to the corresponding expectation. When j is a parent of i, i.e. $i \sim j$ with j on the path between ϕ and i, for notational simplicity we write $m^{\nu}_{i\to\phi}$ instead of $m^{\nu}_{i\to j}$.
Now we claim that for any extremal measure $\nu_{\mathsf{T}} \notin \{\nu_{+,\mathsf{T}},\ \nu_{-,\mathsf{T}}\}$, $\nu_{-,\mathsf{T}}(x_\phi = 1) < \nu_{\mathsf{T}}(x_\phi = 1) < \nu_{+,\mathsf{T}}(x_\phi = 1)$.
By Griffith's inequality we already have $\nu_{\mathsf{T}}(x_\phi = 1) \le \nu_{+,\mathsf{T}}(x_\phi = 1)$ and to prove that the inequality is strict we first note that for any extremal measure ν,

$$\mathrm{atanh}[\nu_{\mathsf{T}}\langle x_\phi\rangle] = \sum_{i\in\partial\phi}\mathrm{atanh}[\tanh(\beta)m^{\nu}_{i\to\phi}] \tag{3.6}$$

Using Griffith's inequality $m^\nu_{i\to\phi} \le m^+_{i\to\phi}$. Thus from (3.6), $\nu_\mathsf{T}\langle x_\phi\rangle = \nu_{+,\mathsf{T}}\langle x_\phi\rangle$ implies that $m^\nu_{i\to\phi} = m^+_{i\to\phi}$ for all $i \in \partial\phi$. Then by induction it follows that $m^\nu_{i\to\phi} = m^+_{i\to\phi}$ for all i. By a similar argument any finite dimensional marginal, at ϕ, under ν_T (or $\nu_{+,\mathsf{T}}$) is determined the values of $\{m^\nu_{i\to\phi}\}$ (or $\{m^+_{i\to\phi}\}$ respectively), hence such finite dimensional marginals coincide under ν_T and $\nu_{+,\mathsf{T}}$, and by Kolmogorov extension theorem $\nu_\mathsf{T} = \nu_{+,\mathsf{T}}$. Therefore we must have strict inequality. Similarly $\nu_{-,\mathsf{T}}(x_\phi = 1) < \nu_\mathsf{T}(x_\phi = 1)$.

Since $\mathsf{T} \in A_2$, pick an extremal Ising Gibbs measure $\nu_\mathsf{T} \notin \{\nu_{+,\mathsf{T}}, \nu_{-,\mathsf{T}}\}$ such that

$$\sum_{i\in\partial\phi} \nu_\mathsf{T}\langle x_\phi \cdot x_i\rangle = \sum_{i\in\partial\phi} \nu_{+,\mathsf{T}}\langle x_\phi \cdot x_i\rangle \tag{3.7}$$

We will show that there exists *only one* $i \in \partial\phi$ such that $m^+_{i\to\phi} > 0$ and $m^+_{\phi\to i} = 0$. Call this property $\mathscr{P}$.

Since $\nu_\mathsf{T} \notin \{\nu_{+,\mathsf{T}}, \nu_{-,\mathsf{T}}\}$. Then there exists at least one $i \in \partial\phi$ such that $m^\nu_{i\to\phi} < m^+_{i\to\phi}$ (by (3.6)). Fix that i.

For any extremal measure ν_T, an easy tree calculation yields,

$$\nu_\mathsf{T}\langle x_\phi \cdot x_i\rangle = F(m^\nu_{\phi\to i} m^\nu_{i\to\phi}) := \frac{\tanh(\beta) + m^\nu_{\phi\to i} m^\nu_{i\to\phi}}{1 + \tanh(\beta) m^\nu_{\phi\to i} m^\nu_{i\to\phi}}.$$

By (3.7) $\nu_\mathsf{T}\langle x_\phi \cdot x_i\rangle = \nu_{+,\mathsf{T}}\langle x_\phi \cdot x_i\rangle$. Now noting that $r \mapsto \frac{a+r}{1+ar}$ is strictly increasing for $0 < a < 1$, and $m^-_{\phi\to i} \le m^\nu_{\phi\to i} \le m^+_{\phi\to i}$,

$$\nu_\mathsf{T}\langle x_\phi \cdot x_i\rangle = \nu_{\mathsf{T},+}\langle x_\phi \cdot x_i\rangle \implies m^+_{i\to\phi} > 0 \text{ and } m^+_{\phi\to i} = 0. \tag{3.8}$$

Since $m^+_{\phi\to i} = 0$ we further have $m^+_{j\to\phi} = 0$ for all $j \in \partial\phi\backslash\{i\}$. Since $m^+_{i\to\phi} > 0$, we further have $m^+_{\phi\to j} > 0$ for all $j \in \partial\phi\backslash\{i\}$, and hence $A_2 \subseteq \{(\mathsf{T}, \phi) : (\mathsf{T}, \phi) \text{ satisfies } \mathscr{P}\} := \widetilde{A}_2$. Thus by assumption $\mu_0(\widetilde{A}_2) \ge \mu_0(A_2) = 1$.

Now for $i \in \partial\phi$, define $f(\phi, i) = \mathbb{I}(m^+_{\phi\to i} = 0)$. Thus by unimodularity

$$\mathbb{E}_{\mu_0}\Big[\sum_{i\in\partial\phi} \mathbb{I}(m^+_{\phi\to i} = 0)\Big] = \mathbb{E}_{\mu_0}\Big[\sum_{i\in\partial\phi} \mathbb{I}(m^+_{i\to\phi} = 0)\Big] \tag{3.9}$$

Since μ_0 supported on $\widetilde{A}_2$, RHS $= \Delta_\phi - 1$ and LHS $= 1$ for μ_0 a.e. every T. Therfore we get $\mathbb{E}_{\mu_0}[\Delta_\phi] = 2$.

Recall that for any $\mu \in \mathcal{U}$, supported on infinite trees, $\overline{deg}(\mu) = 2 \iff \mu$ a.e. T has at most two ends [4, Theorem 6.2]. Thus on μ_0 a.e. every T has at most two ends. Further noting that on any tree with at most two ends, the discrete time SRW is recurrent, we have $\operatorname{br}\mathsf{T} \le 1$ and thus there must be one Ising Gibbs measure on that tree. This suggests for μ_0 a.e. T we must have only one Ising Gibbs measure on T. But this is not possible, as μ_0 is supported on $\widetilde{A}_2$, and from the definition of $\widetilde{A}_2$, for every $\mathsf{T} \in \widetilde{A}_2$, $m^+_\phi > 0$. So contradiction and hence $\mu_0(\widetilde{A}_2) < 1$. $\square$

Lemma 3.5. *For any subsequential limit $\overline{\nu}$, there exists $\alpha_\mathsf{T} \in [0,1]$ Borel measurable in T, such that*

$$\overline{\nu}_\mathsf{T} = \alpha_\mathsf{T}\nu_{+,\mathsf{T}} + (1-\alpha_\mathsf{T})\nu_{-,\mathsf{T}} \text{ for a.s. every } \mathsf{T}. \tag{3.10}$$

Moreover any local subsequential weak limit, $\mathfrak{m}$, is supported on $\mathcal{M}_0 := \{\delta_\mathsf{T}\otimes, \overline{\nu}_\mathsf{T}, \overline{\nu}_\mathsf{T}$ satisfies (3.10)$\}$. Same conclusion holds for $\overline{\nu}_+$ and $\mathfrak{m}_+$, under the uniform bounded degree assumption.

Proof: Setting $\alpha_{\mathsf{T}} = 1/2$ when $\nu_{+,\mathsf{T}} = \nu_{-,\mathsf{T}}$, the result follows from by an application of Lemma 3.1, Lemma 2.3 and Lemma 3.4. Measurability of α_{T} follows from the measurability of $\overline{\nu}_{\mathsf{T}}$, $\overline{\nu}_{+,\mathsf{T}}, \nu_{+,\mathsf{T}}$ and $\nu_{-,\mathsf{T}}$ (in $\mathcal{G}_*$). Upon using Lemma 3.6 (b) we further obtain $\mathfrak{m}$ is supported on $\mathcal{M}_0$. $\square$

Before we provide the proof of the main theorem, we need one more Lemma.

Lemma 3.6. *If $G_n \stackrel{\text{LWC}}{\Longrightarrow} \mu$, and μ_n converge locally to $\mathfrak{m}$, then*
(a) $\mathfrak{m}$ is supported on $\mathcal{M} := \{\delta_{\mathsf{T}} \otimes \nu_{\mathsf{T}},\ \mathsf{T} \in \mathcal{G}_*,\ \nu_{\mathsf{T}} \in \mathcal{P}(\{-1,1\}^{\mathsf{T}}\}$.
(b) Let $\widetilde{\mathfrak{m}}$ be the marginal of $\mathfrak{m}$, a probability measure on $\mathcal{P}(\mathcal{G}_*)$. Then $\widetilde{\mathfrak{m}} = \mu \circ \widetilde{\psi}$, where $\widetilde{\psi}(\mathsf{T}) = \delta_{\mathsf{T}}$, for $\mathsf{T} \in \mathcal{G}_*$. Similar conclusions hold for $\mathfrak{m}_+$.
(c)

$$\lim_{n\to\infty} \mathbb{E}_n\{|\mu_n\langle x_{I_n}\rangle|\} = \mathbb{E}_\mu\Big[\int |\nu_{\mathsf{T}}\langle x_\phi\rangle|\mathfrak{m}(d\nu_{\mathsf{T}})\Big] = 0. \tag{3.11}$$

Proof: Since $\widetilde{\mathfrak{m}}$ is the marginal of $\mathfrak{m}$, to prove (i) we will prove that $\widetilde{\mathfrak{m}}$ is supported on $\{\delta_{\mathsf{T}},\ \mathsf{T} \in \mathcal{G}_*\}$. To this end note that $\mathbb{P}_n^t(I_n)$ is supported on $\{\delta_{B_i(t)} \otimes \mu_{n,B_i(t)}\}$ and since G_n's are locally tree-like it can be easily argued that $\mathfrak{m}^t$ is supported on $\{\delta_{\mathsf{T}(t)} \otimes \nu_{\mathsf{T}(t)}, \mathsf{T} \in \mathcal{G}_*, \nu_{\mathsf{T}(t)} \in \mathcal{P}(\{-1,1\}^{\mathsf{T}(t)})\}$. Therefore, $\widetilde{\mathfrak{m}}^t$, marginal of $\mathfrak{m}^t$, is supported on $\{\delta_{\mathsf{T}}(t), \mathsf{T} \in \mathcal{G}_*\}$. Since this is true for every t, we obtain (a).

Since for each t, $\mathbb{P}_n^t(I_n) \Rightarrow \mathfrak{m}^t$, so does $\widetilde{\mathbb{P}}_n^t(I_n)$, the marginal of $\mathbb{P}_n^t(I_n)$ and the limit is $\widetilde{\mathfrak{m}}^t$. Thus for any $\mathsf{T}_0 \in \mathcal{G}_*$,

$$\begin{aligned}\widetilde{\mathfrak{m}}^t\Big(\delta_{\mathsf{T}(t)} : \mathsf{T}(t) \simeq T_0(t)\Big) = \mathfrak{m}^t(\delta_{\mathsf{T}_0(t)}) = \lim_{n\to\infty} \frac{1}{n}\mathbb{I}(\delta_{B_i(t)} = \delta_{\mathsf{T}_0(t)}) &= \lim_{n\to\infty} \frac{1}{n}\mathbb{I}(B_i(t) \simeq \mathsf{T}_0(t)) \\ &= \mu\Big(B_{\mathsf{T}_0}\Big(\frac{1}{r+1}\Big)\Big),\end{aligned} \tag{3.12}$$

where the last equality follows from $G_n \stackrel{\text{LWC}}{\Longrightarrow} \mu$. Since the above is true for any t and for any $\mathsf{T}_0 \in \mathcal{G}_*$ we prove (b). Proof for $\mathfrak{m}_+$ is similar, we leave the details.

Now to prove (c) note that $\delta_G \otimes \nu_G \stackrel{F}{\mapsto} |\nu_G\langle x_\phi\rangle|$ is continuous, for $G \in \mathcal{G}_*(t),\ \nu_G \in \mathcal{P}(\{-1,1\}^G)$. Thus for any $t \geq 2$,

$$\lim_{n\to\infty} \mathbb{E}_n\{|\mu_n\langle x_{I_n}\rangle|\} = \int F(\overline{\nu}^t)\mathfrak{m}^t(d\overline{\nu}^t).$$

Since $\mathfrak{m}$ is supported on $\mathcal{M}$ and F is a local function, it is easy to see that for every $t \geq 2$, $F(\overline{\nu}^t) = F(\overline{\nu})$ and thus,

$$\lim_{n\to\infty} \mathbb{E}_n\{|\mu_n\langle x_{I_n}\rangle|\} = \int F(\overline{\nu})\mathfrak{m}(d\overline{\nu}).$$

Now further noting that $\widetilde{\mathfrak{m}} = \mu \circ \widetilde{\psi}$, we obtain (3.11). The rightmost equality in (3.11) is trivial, since for each $i \in [n]$, $\mu_n\langle x_i\rangle = 0$. $\square$

Proof of Theorem 1.5: Using Lemma 2.4 obtain a subsequential limit $\mathfrak{m}$. Since $\overline{\nu} \stackrel{F}{\mapsto} \overline{\nu}\Big(\sum_{i\in\partial\phi} x_\phi \cdot x_i\Big)$ is continuous for $\overline{\nu} \in \mathcal{P}(\overline{\mathcal{G}}_*)$, we have

$$\lim_{m\to\infty} \frac{1}{n_m} \sum_{(i,j)\in E_{n_m}} \mu_{n_m}\langle x_i \cdot x_j\rangle = \int \overline{\nu}^t\Big(\sum_{i\in\partial\phi} x_\phi \cdot x_i\Big)\mathfrak{m}^t(d\overline{\nu}^t).$$

Noting F is a local function we further have,

$$\lim_{m\to\infty} \frac{1}{n_m} \sum_{(i,j)\in E_{n_m}} \mu_{n_m}\langle x_i \cdot x_j\rangle = \int \overline{\nu}\Big(\sum_{i\in\partial\phi} x_\phi \cdot x_i\Big)\mathfrak{m}(d\overline{\nu}). \tag{3.13}$$

Using Lemma 3.1, and Lemma 3.6 (b),

$$\mathbb{E}_\mu\Big[\sum_{i\in\partial\phi}\nu_{+,\mathsf{T}}\langle x_\phi\cdot x_i\rangle\Big]=\mathbb{E}_\mu\Big[\int\sum_{i\in\partial\phi}\nu_{\mathsf{T}}\langle x_\phi\cdot x_i\rangle m(d\nu_{\mathsf{T}})\Big]. \tag{3.14}$$

Thus from Lemma 3.4 we further have $\nu_{\mathsf{T}}=\alpha_{\mathsf{T}}\nu_{+,\mathsf{T}}+(1-\alpha_{\mathsf{T}})\nu_{-,\mathsf{T}}$ for a.e. every T and a.e. every ν_{T}, under $\mathfrak{m}$. Therefore $|\nu_{\mathsf{T}}\langle x_\phi\rangle|=|(2\alpha_{\mathsf{T}}-1)|\nu_{+,\mathsf{T}}\langle x_\phi\rangle$ and hence by Lemma 3.6 we conclude that $\alpha_{\mathsf{T}}=1/2$ a.s. under $\mathfrak{m}$. □

4. Proof of Theorem 1.7

In the remaining part of this paper we will argue that any arbitrary values of α_{T} is not possible. To do this we will use some choices of functionals which we will define. Before doing that we need to introduce the following notations:

For any Graph G either finite or infinite, consider continuous time SRW on G. For any $i,j\in V(G)$, denote $a_{i,j}^{l,G}$ to be the average occupation time at vertex j during the time interval $[0,l]$ of the continuous time SRW that starts at i. So formally,

$$a_{i,j}^{l,G}=\frac{1}{l}\int_0^l\mathbb{P}_i(X_t=j)\;dt.$$

When the underlying graph is G_n the average occupation time will be denoted by $a_{i,j}^{l,n}$ and for trees it will be $a_{i,j}^{l,\mathsf{T}}$.

Now for every vertex $i\in V_n$ and $\eta\geq 0$, define

$$F_i^l(\underline{x},\eta):=\mathbb{I}\Big\{\sum_j x_j a_{i,j}^{l,n}\leq -\eta/2\Big\} \tag{4.1}$$

and

$$A_i^{n,l}(\eta):=\mathbb{I}\Big\{\Big|\sum_j\mu_{n,+}\langle x_j\rangle a_{i,j}^{l,n}\Big|\geq\eta\Big\}. \tag{4.2}$$

Similarly as above one can also define $F_i^l(\underline{x},\eta)$ and $A_i^{\mathsf{T},l}(\eta)$ for every vertex $i\in V(\mathsf{T})$, replacing $a_{i,j}^{l,n}$ and $\mu_{n,+}\langle x_j\rangle$ with $a_{i,j}^{l,\mathsf{T}}$ and $\overline{\nu}_{+,\mathsf{T}}\langle x_j\rangle$ respectively. For notational ease we will write F_i^l, $A_i^{n,l}$ and $A_i^{\mathsf{T},l}$ instead of $F_i^l(\underline{x},\eta)$, $A_i^{n,l}(\eta)$ and $A_i^{\mathsf{T},l}(\eta)$ when η and $\underline{x}$ are clear from the context. Sometimes we will also write $F_i^{l,G}$ when we need to be explicit about the underlying graph G.

Lemma 4.1. *Let $G_n\overset{\text{LWC}}{\Longrightarrow}\mu$, with degrees of $\{G_n\}_{n\in\mathbb{N}}$ being uniformly bounded by some finite number Δ. Also assume $\mu_{n,+}$ converges locally weakly to $\mathfrak{m}_+$, and thus in average to $\overline{\nu}_+$, for some $\overline{\nu}_+\in\mathcal{P}(\overline{\mathcal{G}}_*)$. Then for any fixed l and except for a countable many $\eta>0$,*

(a) $\lim_{n\to\infty}\mathbb{E}_n\{\mu_{n,+}(F_{I_n}^l)\}=\overline{\nu}_+(F_\phi^l)$.

(b) $\lim_{n\to\infty}\mathbb{E}_n\Big[\mu_{n,+}\Big\{\sum_{i\in\partial I_n}\mathbb{I}(F_{I_n}^l\neq F_i^l)\Big\}\Big]=\overline{\nu}_+\Big[\Big\{\sum_{i\in\partial\phi}\mathbb{I}(F_\phi^l\neq F_i^l)\Big\}\Big]$

(c) $\lim_{n\to\infty}\mathbb{E}_n\Big[\mu_{n,+}(F_{I_n}^lA_{I_n}^{n,l})\Big]=\mathfrak{m}_+\Big[\overline{\nu}_{+,\mathsf{T}}(F_\phi^lA_\phi^{\mathsf{T},l})\Big]$.

(d) $\lim_{n\to\infty}\mathbb{E}_n\Big[\mu_{n,+}\Big\{\sum_{i\in\partial I_n}\mathbb{I}(F_{I_n}^lA_{I_n}^{n,l}\neq F_i^lA_i^{n,l})\Big\}\Big]=\mathfrak{m}_+\Big[\overline{\nu}_{+,\mathsf{T}}\Big\{\sum_{i\in\partial\phi}\mathbb{I}(F_\phi^lA_\phi^{\mathsf{T},l}\neq F_i^lA_i^{\mathsf{T},l})\Big\}\Big]$.

Remark 4.2. If the functions above were local functions, the conclusions would have followed straight from local weak convergence of $\{\mu_{n_m,+}\}$ and $\{\mu_{n_m}\}$. In this Lemma we first show that these functions are well approximated by some local functions, and then the conclusion follows.

Proof of Lemma 4.1: To prove (a) first note that for any graph G, either finite or infinite, whose degrees are uniformly bounded, for every $i \in V(G)$,

$$\begin{aligned} \sum_{j \notin B_i(M)} a_{i,j}^{l,G} &= \mathbb{E}_i\Big[\frac{1}{l}\int_0^l \mathbb{I}(X_t \notin B_i(M))dt\Big] \\ &\leq \mathbb{P}_i\Big\{\text{there are at least } M \text{ transitions in } [0,l]\Big\} \\ &\leq \mathbb{P}\Big[\sum_{j=1}^{M} E_j \leq l\Big] \\ &= \sum_{i=M}^{\infty} \frac{(\Delta l)^i}{i!} e^{-\Delta l} \leq \frac{M}{M-\Delta l} e^{-\Delta l}\frac{(\Delta l)^M}{M!} \equiv \varepsilon_M/2, \end{aligned} \tag{4.3}$$

where $M \equiv M(l)$ chosen such that $M > \Delta l$ and $E_i \overset{i.i.d.}{\sim} Exp(\Delta)$.
Further define,

$$F_i^{l,M,G}(\eta) \equiv F_i^{l,M,G}(\underline{x},\eta) = \mathbb{I}\Big(\sum_{j \in B_i(M)} x_j a_{i,j}^{l,G} \leq -\eta/2\Big).$$

Thus by (4.3) for any graph G and every $i \in V(G)$,

$$F_i^{l,M,G}(\eta + \varepsilon_M) \leq F_i^{l,G}(\eta) \leq F_i^{l,M,G}(\eta - \varepsilon_M). \tag{4.4}$$

Further note that if we have two graphs G_1 and G_2 such that $G_1(N) \simeq G_2(N)$, then two markov chains can be trivially coupled, so that they agree up to time l on the event that there are less than N transitions in time $[0,l]$. Thus for any $N > M$,

$$F_i^{l,M,G_2}(\eta + \varepsilon_M + 2\varepsilon_N) \leq F_i^{l,M,G_1}(\eta + \varepsilon_M) \text{ and } F_i^{l,M,G_1}(\eta - \varepsilon_M) \leq F_i^{l,M,G_1}(\eta - \varepsilon_M - 2\varepsilon_N). \tag{4.5}$$

Using $\mathbb{P}^N_{n_m,+} \Rightarrow \overline{\nu}^N_+$, and by (4.4), and (4.5)we get

$$\begin{aligned} \overline{\nu}_+^N\Big(F_\phi^{l,M,T(N)}(\eta + \varepsilon_M + 2\varepsilon_N)\Big) &\leq \liminf_{n\to\infty} \mathbb{E}_n\{\mu_n(F^l_{I_n})\} \\ &\leq \limsup_{n\to\infty} \mathbb{E}_n\{\mu_n(F^l_{I_n})\} \leq \overline{\nu}_+^N\Big(F_\phi^{l,M,T(N)}(\eta - \varepsilon_M - 2\varepsilon_N)\Big) \end{aligned}$$

and

$$\overline{\nu}_+^N\Big(F_\phi^{l,M,T(N)}(\eta + \varepsilon_M + 2\varepsilon_N)\Big) \leq \overline{\nu}_+(F^l_\phi) \leq \overline{\nu}_+^N\Big(F_\phi^{l,M,T(N)}(\eta - \varepsilon_M - 2\varepsilon_N)\Big).$$

Since N, M are arbitrary and $\varepsilon_M,\ \varepsilon_N \to 0$ as $M,\ N \to \infty$, the proof is completed by taking η to be a point of continuity of the distribution of F^l_ϕ under $\overline{\nu}_+$.
Now to prove (b) we first note that

$$\widetilde{F}(\eta) := \overline{\nu}_+\Big[\Big\{\sum_{i \in \partial\phi} \mathbb{I}(F^l_\phi(\eta) = 1, F^l_i(\eta) = 1)\Big\}\Big] \tag{4.6}$$

is a decreasing, left continuous function in η and $\lim_{\eta \to -\infty} \widetilde{F}(\eta) < \infty$. Thus points of discontinuity of $\widetilde{F}(\ \cdot\ )$ is at most countable. Now choosing η to be a point of continuity of $\widetilde{F}$, and a point of continuity of the distribution of F^l_ϕ under $\overline{\nu}_+$, we have,

$$\widetilde{F}_1(\eta) := \overline{\nu}_+\Big[\Big\{\sum_{i \in \partial\phi} \mathbb{I}(F^l_\phi(\eta) = 1, F^l_i(\eta) = 0)\Big\}\Big],$$

is continuous at η. Further by unimodularity of μ,

$$\widetilde{F}_1(\eta) = \overline{\nu}_+\Big[\Big\{\sum_{i\in\partial\phi}\mathbb{I}(F^l_\phi(\eta)=0, F^l_i(\eta)=1)\Big\}\Big].$$

Now proof of (b) is completed by adapting the proof of (a).
Now we prove (c): Since both $\mathbb{P}^t_n(I_n)$ and $\mathfrak{m}^t_+$ are supported on $\{\delta_G\otimes\nu_G, G\in\mathcal{G}_*(t), \nu_G\in\mathcal{P}(\{-1,1\}^G)\}$, noting that

$$\delta_G\otimes\nu_G \overset{\Psi}{\longmapsto} \nu_G\Big[F_1\Big(\sum_{j\in B_\phi(G)} x_j a^{l,G}_{\phi,j}\Big)\Big]F_2\Big(\sum_{j\in B_\phi(G)}\nu_G\langle x_j\rangle a^{l,G}_{\phi,j}\Big),$$

is continuous whenever $F_1, F_2:\mathbb{R}\mapsto\mathbb{R}$ are continuous functions, from local weak convergence of $\mu_{n,+}$ it follows that $\mathbb{P}^t_n(I_n)[\Psi]\to\mathfrak{m}^t_+[\Psi]$. Furthermore

$$\widetilde{F}_2(\eta) := \mathfrak{m}_+\Big(F^l_\phi(\eta)A^{\mathsf{T},l}_\phi(\eta)\Big) \tag{4.7}$$

is decreasing, left-continuous function of η, with finite limits at $\pm\infty$ and hence discontinuous for at most countably many choices of η. Thus noting that, for any η there exists F_+, F_- continuous such that $\mathbb{I}(x\le\eta-\varepsilon)\le F_-(x)\le\mathbb{I}(x\le\eta)\le F_+(x)\le\mathbb{I}(x\le\eta+\varepsilon)$, the proof follows by adapting the proof of (a). Proof of (d) is similar and hence omitted. $\square$

Lemma 4.3. *Fix any $\eta>0$ and fix a configuration $\underline{x}=\underline{x}(n)$ belonging to the support of $\mu_{+,n}$. Let U_n denote the uniform measure on the vertices of G_n. Then for every n and l,*

$$\mathbb{E}_{U_n}(F^l_{I_n}) \le \frac{1}{1+\eta/2} \tag{4.8}$$

Proof of Lemma 4.3: First note that for any vertex i,

$$\sum_j a^{l,n}_{i,j} = 1$$

and for any two vertices i and j, $a^{l,n}_{i,j}=a^{l,n}_{j,i}$, since the markov chain under consideration is reversible. Now noting that

$$\mathbb{E}_{U_n}(\sum_j x_j a^{l,n}_{I_n,j}) = \frac{1}{n}\sum_j x_j\sum_i a^{l,n}_{i,j} \ge 0$$

and $|\sum_j x_j a^{l,n}_{i,j}|\le 1$ for every i, we complete the proof by an use of Markov's inequality. $\square$

Lemma 4.4. *For $\mathsf{T}\in\mathcal{G}_*$ consider any Ising Gibbs measure $\overline{\nu}_{+,\mathsf{T}}$ on T. For every $i\in V(\mathsf{T})$, let $\overline{\rho}^l_{i,\mathsf{T}}=\sum_j\overline{\nu}_{+,\mathsf{T}}\langle x_j\rangle a^{l,\mathsf{T}}_{i,j}$. Then for any $\mu\in\mathcal{U}$, and any fixed $\varepsilon>0$,*

$$\mathbb{E}_\mu\Big[\sum_{i\in\partial\phi}\mathbb{I}\Big(|\overline{\rho}^l_{\phi,\mathsf{T}}-\overline{\rho}^l_{i,\mathsf{T}}|>\varepsilon\Big)\Big]\to 0 \text{ as } l\to\infty. \tag{4.9}$$

Proof: First we will show that (4.9) holds when $\mu^e\in\mathcal{U}$ is an ergodic measure.
First note that

$$\rho := \mathbb{E}_\mu\Big[\sum_j\overline{\nu}_{+,\mathsf{T}}\langle x_j\rangle a^{l,\mathsf{T}}_{\phi,j}\Big] = \mathbb{E}_\mu\Big[\overline{\nu}_{+,\mathsf{T}}\langle x_\phi\rangle\sum_j a^{l,\mathsf{T}}_{\phi,j}\Big] = \mathbb{E}_\mu\Big[\overline{\nu}_{+,\mathsf{T}}\langle x_\phi\rangle\Big], \tag{4.10}$$

where the middle identity follows by unimodularity of μ. Now consider the stationary probability measure σ^e for the SRW, where $\frac{d\sigma^e}{d\mu^e} = \frac{\Delta_\phi}{\overline{deg}(\mu^e)}$, and hence by triangle inequality,

$$\begin{aligned}
\mathbb{E}_{\mu^e}\Big[\sum_{i\in\partial\phi}\mathbb{I}\Big(|\overline{\rho}^l_{\phi,\mathsf{T}}-\overline{\rho}^l_{i,\mathsf{T}}|>\varepsilon\Big)\Big] &\le \mathbb{E}_{\mu^e}\Big[\Delta_\phi\mathbb{I}\Big(|\overline{\rho}^l_{\phi,\mathsf{T}}-\rho|>\varepsilon/2\Big)\Big]+\mathbb{E}_{\mu^e}\Big[\sum_{i\in\partial\phi}\mathbb{I}\Big(|\overline{\rho}^l_{i,\mathsf{T}}-\rho|>\varepsilon/2\Big)\Big]\\
&= \frac{\mathbb{E}_{\sigma^e}\Big[\mathbb{I}\Big(|\overline{\rho}^l_{\phi,\mathsf{T}}-\rho|>\varepsilon/2\Big)\Big]+\mathbb{E}_{\sigma^e}\Big[\frac{1}{\Delta_\phi}\sum_{i\in\partial\phi}\mathbb{I}\Big(|\overline{\rho}^l_{i,\mathsf{T}}-\rho|>\varepsilon/2\Big)\Big]}{\overline{deg}(\mu^e)}\\
&= \frac{2}{\overline{deg}(\mu^e)}\mathbb{E}_{\sigma^e}\Big[\mathbb{I}\Big(|\overline{\rho}^l_{\phi,\mathsf{T}}-\rho|>\varepsilon/2\Big)\Big]. \qquad (4.11)
\end{aligned}$$

Next recall that

$$\overline{\rho}^l_{\phi,\mathsf{T}} = \frac{1}{l}\int_0^l\sum_j\overline{\nu}_{+,\mathsf{T}}\langle x_j\rangle\mathbb{P}_\phi(X_t=j)dt = \mathbb{E}^{\text{SRW}}_\phi\Big[\frac{1}{l}\int_0^l\overline{\nu}_{+,\mathsf{T}}\langle x_{X_t}\rangle dt\Big] \qquad (4.12)$$

Let $(Y_n)_{n\ge0}$ denote the discrete time Markov chain embedded in $t\mapsto X_t$. Then, with

$$N_l := \max\Big\{i:\sum_{j=1}^i\frac{E_j}{\Delta_{Y_{j-1}}}\le l\Big\},\ \Gamma_{N_l}=\sum_{j=1}^{N_l}\frac{E_j}{\Delta_{Y_{j-1}}},$$

and $\{E_i\}$ i.i.d. Exp(1) independent of $(Y_n)_{n\ge0}$, clearly we have, the identity in law,

$$\frac{1}{l}\int_0^l\overline{\nu}_{+,\mathsf{T}}\langle x_{X_t}\rangle dt = \frac{1}{l}\Big[\sum_{j=1}^{N_l}\overline{\nu}_{+,\mathsf{T}}\langle x_{Y_{j-1}}\rangle\frac{E_j}{\Delta_{Y_{j-1}}}\Big]+\frac{l-\Gamma_{N_l}}{l}\overline{\nu}_{+,\mathsf{T}}\langle x_{Y_{N_l}}\rangle. \qquad (4.13)$$

Since $(Y_n)_{n\ge0}$ is ergodic for its stationary initial distribution σ^e, by Birkhoff's ergodic theorem have that, for σ^e-a.e.

$$\frac{1}{l}\sum_{j=1}^l\frac{1}{\Delta_{Y_{j-1}}}\to\mathbb{E}_{\sigma^e}\Big[\frac{1}{\Delta_{Y_\phi}}\Big]=\frac{1}{\overline{deg}(\mu^e)}. \qquad (4.14)$$

This obviously implies that (see [23, Corollary 3.22]),

$$\frac{1}{l}\sum_{j=1}^l\frac{E_j}{\Delta_{Y_{j-1}}}\xrightarrow{\text{a.s.}}\frac{1}{\overline{deg}(\mu^e)} \qquad (4.15)$$

Noting that $N_l\to\infty$ and using (4.15), applying standard renewal theorem techniques we get,

$$\frac{N_l}{l}\xrightarrow{\text{a.s.}}\overline{deg}(\mu^e),$$

and thus by (4.15),

$$\frac{\Gamma_{N_l}}{l}\xrightarrow{\text{a.s.}}1.$$

Therefore from (4.13), upon using ergodic theorem, and [23, Corollary 3.22],

$$\overline{\rho}^l_{\phi,\mathsf{T}}=\frac{1}{l}\int_0^l\overline{\nu}_{+,\mathsf{T}}\langle x_{X_t}\rangle dt\xrightarrow{\text{a.s.}}\overline{deg}(\mu^e)\times\mathbb{E}_{\sigma^e}\Big[\frac{\overline{\nu}_{+,\mathsf{T}}\langle x_\phi\rangle}{\Delta_\phi}\Big]=\mathbb{E}_{\mu^e}\Big[\overline{\nu}_{+,\mathsf{T}}\langle x_\phi\rangle\Big]=\rho. \qquad (4.16)$$

Hence we get (4.9) for ergodic measures, from (4.11) and (4.12).
Since every $\mu\in\mathcal{U}$ can be written as a *Choquet integral* of extremal measures [4, Lemma 6.8], and any extremal measure is ergodic, we have the result for all $\mu\in\mathcal{U}$. $\square$

Lemma 4.5. *For any $\mathsf{T} \in \mathcal{G}_*$ consider any Ising Gibbs measure $\overline{\nu}_{+,\mathsf{T}}$ such that $\overline{\nu}_{+,\mathsf{T}} = \alpha_{\mathsf{T}}\nu_{+,\mathsf{T}} + (1-\alpha_{\mathsf{T}})\nu_{-,\mathsf{T}}$ for some Borel measurable $\alpha_{\mathsf{T}} \in [0,1]$. Then for any $\eta > 0$, except possibly a countably many, under probability measure $\mu \in \mathcal{U}$,*
(a) $\nu_{+,\mathsf{T}}\Big(F^l_\phi = 1\Big)\mathbb{I}[A^{\mathsf{T},l}_\phi = 1] \stackrel{\mathcal{P}}{\to} 0$ *as* $l \to \infty$.
(b) $\nu_{-,\mathsf{T}}\Big(F^l_\phi = 0\Big)\mathbb{I}[A^{\mathsf{T},l}_\phi = 1] \stackrel{\mathcal{P}}{\to} 0$ *as* $l \to \infty$.
(c) $\sum_{i\in\partial\phi} \nu_{+,\mathsf{T}}\Big(F^l_\phi A^{\mathsf{T},l}_\phi \neq F^l_i A^{\mathsf{T},l}_i\Big) \stackrel{\mathcal{P}}{\to} 0$ *as* $l \to \infty$.
(d) $\sum_{i\in\partial\phi} \nu_{-,\mathsf{T}}\Big(F^l_\phi A^{\mathsf{T},l}_\phi \neq F^l_i A^{\mathsf{T},l}_i\Big) \stackrel{\mathcal{P}}{\to} 0$ *as* $l \to \infty$.

Proof of Lemma 4.5: We will prove (b) and (d). Proof of other parts are similar and hence details are omitted.
Since $0 \le \nu_{-,\mathsf{T}}\Big(F^l_\phi = 0\Big) \le 1$ to prove (b) it is enough to prove

$$\mathbb{E}_\mu\Big[\nu_{-,\mathsf{T}}\Big(F^l_\phi = 0\Big)A^{\mathsf{T},l}_\phi\Big] \to 0 \text{ as } l \to \infty.$$

Note that

$$\mathbb{E}_\mu\Big[\nu_{-,\mathsf{T}}\Big(F^l_\phi = 0\Big)A^{\mathsf{T},l}_\phi\Big] = \mathbb{E}_\mu\Big[\nu_{-,\mathsf{T}}\Big\{\sum_i x_i a^{l,\mathsf{T}}_{\phi,i} - \rho^l_{-,\mathsf{T}} > -(\eta/2 + \rho^l_{-,\mathsf{T}})\Big\}A^{\mathsf{T},l}_\phi\Big], \tag{4.17}$$

where $\rho^l_{-,\mathsf{T}} := \sum_j \nu_{-,\mathsf{T}}\langle x_j\rangle a^{l,\mathsf{T}}_{\phi,j}$. Since

$$\Big\{A^{\mathsf{T},l}_\phi = 1\Big\} = \Big\{|\alpha_{\mathsf{T}} - \beta_{\mathsf{T}}|\sum_j \nu_{+,\mathsf{T}}\langle x_j\rangle a^{l,\mathsf{T}}_{\phi,j} \ge \eta\Big\} \subseteq \Big\{\rho^l_{-,\mathsf{T}} \le -\eta\Big\},$$

we get from (4.17), upon using Markov inequality,

$$\begin{aligned}\mathbb{E}_\mu\Big[\nu_{-,\mathsf{T}}\Big(F^l_\phi = 0\Big)A^{\mathsf{T},l}_\phi\Big] &\le \mathbb{E}_\mu\Big[\nu_{-,\mathsf{T}}\Big\{\sum_i x_i a^{l,\mathsf{T}}_{\phi,i} - \rho^l_{-,\mathsf{T}} > \eta/2\Big\}A^{\mathsf{T},l}_\phi\Big] \\ &\le \frac{4}{\eta^2}\mathbb{E}_\mu\Big[\mathrm{Var}_{\nu_{-,\mathsf{T}}}\Big(\sum_i x_i a^{l,\mathsf{T}}_{\phi,i}\Big)A^{\mathsf{T},l}_\phi\Big]\end{aligned}$$

By reversibility of SRW $a^{l,\mathsf{T}}_{\phi,i} = a^{l,\mathsf{T}}_{i,\phi}$; the fact that the covariance of x_i and x_j under plus and minus measure are same; hence expanding the variance of the sum, by unimodularity

$$\mathbb{E}_\mu\Big[\mathrm{Var}_{\nu_{-,\mathsf{T}}}\Big(\sum_i x_i a^{l,\mathsf{T}}_{\phi,i}\Big)A^{\mathsf{T},l}_\phi\Big] = \mathbb{E}_\mu\Big[\sum_j \mathrm{Cov}_{\nu_{+,\mathsf{T}}}(x_\phi, x_j)\sum_i a^{l,\mathsf{T}}_{\phi,i}a^{l,\mathsf{T}}_{i,j}A^{\mathsf{T},l}_i\Big]$$

Now to control the above expression we break the sum over j in the RHS into two parts: In Term I, sum is taken over $j \in B_\phi(r)$, and in Term II, sum is taken over $j \notin B_\phi(r)$, for some fixed positive integer r.

Term II will be controlled by controlling the correlation decay. We will prove that

$$0 \le \mathrm{Cov}_{\nu_{+,\mathsf{T}}}(x_\phi, x_j) \le A\gamma^{d(\phi,j)}, \gamma \in (0,1), \tag{4.18}$$

where the constants A and γ depend on β but independent of the choice of the tree T.
To this end note that on a T, the restriction of $\nu_{+,\mathsf{T}}$ to any fixed path $(x_{v_0}, x_{v_1}, \ldots, x_{v_k})$ is a Markov chain of state space $\{-1,1\}$. For non-regular T this chain is typically non-homogeneous. Nevertheless in [8, Lemma 4.1] it is shown that for μ, an Ising measure on a finite T, with inverse temperature parameter $\{\beta_e\}_{e\in E(T)}$ and external magnetic field parameter $\{B_v\}_{v\in V(T)}$, for any two

vertices $v \neq w$, the value $\mu[x_w|x_v = 1] - \mu[x_w|x_v = -1]$ is maximized when the external magnetic field $B \equiv 0$. Since $x_v \in \{-1, 1\}$, we get

$$\mathrm{Cov}_\mu(x_v, x_w) = 2\mu[x_v = 1]\mu[x_v = -1](\mu[x_w|x_v = 1] - \mu[x_w|x_v = -1]) \tag{4.19}$$

Thus, to bound the covariance between x_v and x_w for any two vertices $v,\ w$ one can assume that external magnetic field parameters corresponding to the marginal Ising measure on the path joining v and w, are identically 0, and this ensures that the Markov chain is homogeneous. Hence using correlation decay formula for 2-state non-degenerate (non-degeneracy here follows from the finiteness of β) homogeneous markov chain we obtain the upper bound in (4.18). Since for any *ferromagnetic* Ising measure μ on a finite T and any two vertices $v,\ w$, $\mu[x_w|x_v = 1] - \mu[x_w|x_v = -1] \geq 0$, lower bound in (4.18) follows from (4.19).
Now by non-negativity of $a_{i,j}^{l,\mathsf{T}}$ and the fact $\sum_{i,j} a_{\phi,i}^{l,\mathsf{T}} a_{i,j}^{l,\mathsf{T}} = 1$, upon using (4.18)

$$\text{Term II} \leq A\left[\sum_{k=r+1}^{\infty} \gamma^k \mathbb{E}_\mu\Big(\sum_{j \in \partial B_\phi(k)} \sum_i a_{\phi,i}^{l,\mathsf{T}} a_{i,j}^{l,\mathsf{T}}\Big)\right] \leq A\gamma^r \mathbb{E}_\mu\Big(\sum_{j \notin B_\phi(r)} \sum_i a_{\phi,i}^{l,\mathsf{T}} a_{i,j}^{l,\mathsf{T}}\Big) \leq A\gamma^r \tag{4.20}$$

Now we control Term I: First note that on any infinite tree T, if $\nu_{+,\mathsf{T}} = \nu_{-,\mathsf{T}}$, then $A_i^{\mathsf{T},l} = 0$ for any $i \in V(\mathsf{T})$ and any positive integer l. It is known that for any infinite tree, the Ising model with no external magnetic field, phase transition occurs at β_c, where β_c is the solution of the equation $[\mathrm{br}\,\mathsf{T}] \times \tanh(\beta) = 1$ (e.g. see [27, Theorem 1.1]). Thus for those tree which have $\mathrm{br}\,\mathsf{T} \leq 1$, there is no phase transition, whereby Term I is 0. So without loss of generality assume hereafter $[\mathrm{br}\,\mathsf{T}] > 1$ and note that

$$\begin{aligned}\sum_{j \in B_\phi(r)} \sum_i a_{\phi,i}^{l,\mathsf{T}} a_{i,j}^{l,\mathsf{T}} &= \sum_{j \in B_\phi(r)} \frac{1}{l^2} \int_0^l \int_0^l \sum_i \mathbb{P}_\phi(X_t = i)\mathbb{P}_i(X_s = j)\ dt\ ds \\ &= \frac{1}{l^2} \int_0^l \int_0^l \mathbb{P}_\phi(X_{t+s} \in B_\phi(r))\ dt\ ds. \end{aligned} \tag{4.21}$$

Since $\mathrm{br}\,\mathsf{T} > 1$, the discrete time SRW on T is transient (see [28, Theorem 4.3]). Hence $\{X_t\}_{t \geq 0}$ is transient and in particular $1 \geq \mathbb{P}_\phi(X_t \in B_\phi(r)) \to 0$ as $t \to \infty$ for any fixed positive integer r. Thus Term I goes to 0 as $l \to \infty$, completing the proof of (b).

Now we prove (d): Since $F_\phi^l A_\phi^{l,\mathsf{T}}, F_i^l A_i^{l,\mathsf{T}} \in \{0, 1\}$, and by unimodularity

$$\Lambda_l := \mathbb{E}_\mu\Big[\sum_{i \in \partial\phi} \nu_{-,\mathsf{T}}\Big\{F_\phi^l A_\phi^{\mathsf{T},l} = 0, F_i^l A_i^{\mathsf{T},l} = 1\Big\}\Big] = \mathbb{E}_\mu\Big[\sum_{i \in \partial\phi} \nu_{-,\mathsf{T}}\Big\{F_\phi^l A_\phi^{\mathsf{T},l} = 1, F_i^l A_i^{\mathsf{T},l} = 0\Big\}\Big],$$

it suffices to show that the LHS of the above expression goes to 0. To this end note that

$$\Big\{F_\phi^l A_\phi^{\mathsf{T},l} = 0, F_i^l A_i^{\mathsf{T},l} = 1\Big\} \subseteq \Big\{F_\phi^l = 1, A_\phi^{l,\mathsf{T}} = 1\Big\} \cup \Big\{A_\phi^{l,\mathsf{T}} = 0, A_i^{l,\mathsf{T}} = 1\Big\}. \tag{4.22}$$

Thus

$$\Lambda_l \leq \mathbb{E}_\mu\Big[\Delta_\phi \nu_{-,\mathsf{T}}\Big(F_\phi^l = 0\Big) A_\phi^{l,\mathsf{T}}\Big] + \mathbb{E}_\mu\Big[\sum_{i \in \partial\phi} \mathbb{I}(A_\phi^{l,\mathsf{T}} = 0, A_i^{l,\mathsf{T}} = 1)\Big]. \tag{4.23}$$

Now the first term in the RHS of the above expression goes to 0 by an use of (b) and DCT. Therefore it only remains to prove

$$\mathbb{E}_\mu\Big[\sum_{i \in \partial\phi} \mathbb{I}(A_\phi^{l,\mathsf{T}} = 0, A_i^{l,\mathsf{T}} = 1)\Big] \to 0 \text{ as } l \to \infty. \tag{4.24}$$

To this end note that

$$\sum_{i\in\partial\phi}\mathbb{I}\Big(A_\phi^{l,\mathsf{T}}=0, A_i^{l,\mathsf{T}}=1\Big)\le\sum_{i\in\partial\phi}\mathbb{I}\Big(|\overline{\rho}_{\phi,\mathsf{T}}^l-\overline{\rho}_{i,\mathsf{T}}^l|>\varepsilon\Big)+\Delta_\phi\mathbb{I}\Big(|\overline{\rho}_{\phi,\mathsf{T}}^l|\in[\eta-\varepsilon,\eta)\Big),$$

Upon using Lemma 4.4 we note that we only need to prove that, for any $\mu\in\mathcal{U}$,

$$\lim_{\varepsilon\to 0}\lim_{l\to\infty}\mathbb{E}_\mu\Big[\mathbb{I}\Big(|\rho_{\phi,\mathsf{T}}^l|\in[\eta-\varepsilon,\eta)\Big)\Big]=0,\tag{4.25}$$

except for a countable many choices of η. First note that for every ergodic measure μ^e, for μ^e a.e. every T, $\overline{\rho}_{\phi,\mathsf{T}}^l\xrightarrow{a.s.}\rho$. Thus using extremal decomposition of $\mu\in\mathcal{U}$, and by DCT, for any set $A\in\mathscr{B}(\mathbb{R})$

$$\theta(A):=\lim_{l\to\infty}\mu\Big(\overline{\rho}_{\phi,\mathsf{T}}^l\in A\Big)=\int\lim_{l\to\infty}\mu^e\Big(\overline{\rho}_{\phi,\mathsf{T}}^l\in A)\Theta(d\mu^e)=\int\mathbb{I}(\rho^e\in A)\Theta(d\mu^e),\tag{4.26}$$

where Θ is a probability measure on $\{\mu^e:\mu^e\in\mathcal{U}\}$ and ρ^e denote the value of ρ (defined in (4.10)) for the measure μ^e. Now it is easy to note that θ is a probability measure on $\mathbb{R}$. Thus choosing η to be continuity point, we obtain the result. $\square$

Proof of Theorem 1.7 (a): We will show that edge-expansion property of $\{G_n\}_{n\in\mathbb{N}}$ implies

$$\mathbb{E}_\mu\Big[\mathbb{I}\Big\{\liminf_{l\to\infty}\sum_j\nu_{+,\mathsf{T}}\langle x_j\rangle a_{\phi,j}^{l,\mathsf{T}}>0\Big\}\int\beta_{\mathsf{T}}\mathbb{I}(\beta_{\mathsf{T}}\neq 1/2)\mathfrak{m}_+(d\overline{\nu}_{+,\mathsf{T}})\Big]=0,\tag{4.27}$$

and will argue that (4.27) further implies the following:

$$\mathbb{E}_\mu\Big[\mathbb{I}\{\nu_{+,\mathsf{T}}\neq\nu_{-,\mathsf{T}}\}\int\beta_{\mathsf{T}}\mathbb{I}(\beta_{\mathsf{T}}\neq 1/2)\mathfrak{m}_+(d\overline{\nu}_{+,\mathsf{T}})\Big]=0.\tag{4.28}$$

Thus we get $\beta_{\mathsf{T}}\in\{0,1/2\}$ for μ-a.e. every T and hence proof will be complete.
To prove (4.28) from (4.27), it is enough to prove that for any extremal measure $\mu^e\in\mathcal{U}$, $\mu^e(\mathscr{S}_\pm\backslash\mathscr{S})=0$, where $\mathscr{S}_\pm=\{\mathsf{T}:\nu_{+,\mathsf{T}}\neq\nu_{-,\mathsf{T}}\}$ and $\mathscr{S}=\{\mathsf{T}:\liminf_{l\to\infty}\sum_j\nu_{+,\mathsf{T}}\langle x_j\rangle a_{\phi,j}^{l,\mathsf{T}}>0\}$. Now adapting the proof of (4.16), we further get that for any $\mu^e\in\mathcal{U}$, $\liminf_{l\to\infty}\sum_j\nu_{+,\mathsf{T}}\langle x_j\rangle a_{\phi,j}^{l,\mathsf{T}}=\mathbb{E}_{\mu^e}[\nu_{+,\mathsf{T}}\langle x_\phi\rangle]$ for μ^e-a.e. every T. Since the limit is constant for μ^e-a.e. every T the desired result is obvious.
So now it only remains to show that edge-expansion property implies (4.27). Suppose not, i.e. there exists an $\varepsilon_0>0$ such that

$$\mathbb{E}_\mu\Big[\mathbb{I}\Big\{\liminf_{l\to\infty}\sum_j\nu_{+,\mathsf{T}}\langle x_j\rangle a_{\phi,j}^{l,\mathsf{T}}>0\Big\}\int\beta_{\mathsf{T}}\mathbb{I}(\beta_{\mathsf{T}}\neq 1/2)\mathfrak{m}_+(d\overline{\nu}_{+,\mathsf{T}})\Big]=\varepsilon_0.$$

Thus there exists $\eta_0(\varepsilon_0)$ such that for all $\eta\le\eta_0(\varepsilon_0)$,

$$\mathbb{E}_\mu\Big[\int\beta_{\mathsf{T}}\mathbb{I}\Big\{|\alpha_{\mathsf{T}}-\beta_{\mathsf{T}}|\sum_j\nu_{+,\mathsf{T}}\langle x_j\rangle a_{\phi,j}^{l,\mathsf{T}}\ge\eta\text{ eventually in }l\Big\}\mathfrak{m}_+(d\overline{\nu}_{+,\mathsf{T}})\Big]\ge\varepsilon_0/2.\tag{4.29}$$

Fix any $\eta>0$ such that Lemma 4.1 and Lemma 4.5 holds. Now for any η and l, let

$$S_{\eta,l}=\{i:F_i^lA_i^{n,l}=1\}$$

Since $\{G_n\}_{n\in\mathbb{N}}$ are $(\delta, 1/2, \lambda_\delta)$ expanders: for $\delta \equiv \frac{\eta}{2+\eta}$, and whenever $\sum_{i=1}^n F_i^l A_i^{n,l} \ge n\delta$ we have

$$\begin{aligned}\sum_{(i,j)\in E_n} \mathbb{I}(F_i^l A_i^{n,l} \neq F_j^l A_j^{n,l}) &\geq \lambda_\delta \min\Big\{\sum_{i\in V_n} F_i^l A_i^{n,l}, \sum_{i\in V_n}(1-F_i^{n,l}A_i^{n,l})\Big\} \\ &\geq \lambda_\delta \min\Big\{\sum_{i\in V_n} F_i^l A_i^{n,l}, \sum_{i\in V_n}(1-F_i^{l})\Big\} \\ &\geq \lambda_\delta \min\Big\{\sum_{i\in V_n} F_i^l A_i^{n,l}, \frac{n\eta}{2+\eta}\Big\} \geq \frac{\lambda_\delta\eta}{2+\eta}\sum_{i\in V_n} F_i^l A_i^{n,l},\end{aligned}$$

where the second last inequality follows from Lemma 4.3. Thus taking expectation on both sides,

$$\mathbb{E}_n\Big[\mu_{n,+}\Big\{\sum_{i\in\partial I_n}\mathbb{I}(F_{I_n}^l A_{I_n}^{n,l} \neq F_i^l A_i^{n,l})\Big\}\Big] \geq \frac{\lambda_\delta\eta}{2+\eta}\mathbb{E}_n\Big[\mu_{n,+}\Big\{F_{I_n}^l A_{I_n}^{n,l}\mathbb{I}\Big(\frac{1}{n}\sum_{i=1}^n F_i^l A_i^{n,l} \geq \delta\Big)\Big\}\Big] \tag{4.30}$$

Now we note that for any non-negative random variable X and any positive real s we have the following inequality,

$$\mathbb{P}(X \geq x) \geq \mathbb{E}\Big[\frac{X-x}{s+X}\Big].$$

Noting that $F_i^l A_i^{n,l} \in \{0,1\}$, and taking $s = 1$ from (4.30) we deduce that

$$\mathbb{E}_n\Big[\mu_{n,+}\Big\{\sum_{i\in\partial I_n}\mathbb{I}(F_{I_n}^l A_{I_n}^{n,l} \neq F_i^l A_i^{n,l})\Big\}\Big] \geq \frac{1}{2}\lambda_\delta\Big(\frac{\eta}{2+\eta}\Big)^2\Big(\mathbb{E}_n\Big[\mu_{n,+}\Big(F_{I_n}^l A_{I_n}^{n,l}\Big)\Big] - \frac{\eta}{2+\eta}\Big).$$

Now from Lemma 2.4 subsequential local weak limit $\mathfrak{m}_+$ exists, for a subsequnce $\{n_m\}$. Thus sending $m\to\infty$, using Lemma 4.1 for $\{\mu_{n_m,+}\}$ we get

$$\mathfrak{m}_+\Big[\overline{\nu}_{+,\mathsf{T}}\Big\{\sum_{i\in\partial I_n}\mathbb{I}(F_{I_n}^l A_{I_n}^{n,l} \neq F_i^l A_i^{n,l})\Big\}\Big] \geq \frac{1}{2}\lambda_\delta\Big(\frac{\eta}{2+\eta}\Big)^2\Big(\mathfrak{m}_+\Big[\overline{\nu}_{+,\mathsf{T}}\Big(F_{I_n}^l A_{I_n}^{n,l}\Big)\Big] - \frac{\eta}{2+\eta}\Big).$$

Further sending $l\to\infty$, from Lemma 4.5 and Lemma 3.5, upon using Lemma 3.6 (b) and Fatou's Lemma we get,

$$0 \geq \frac{1}{2}\lambda_\delta\Big(\frac{\eta}{2+\eta}\Big)^2\Big\{\mathbb{E}_\mu\Big[\int \beta_{\mathsf{T}}\liminf_{l\to\infty}\mathbb{I}(A_\phi^{\mathsf{T},l}=1)\mathfrak{m}_+(d\overline{\nu}_{+,\mathsf{T}})\Big] - \frac{\eta}{2+\eta}\Big\}. \tag{4.31}$$

Since

$$\liminf_{l\to\infty}\mathbb{I}(A_\phi^{\mathsf{T},l}=1) = \mathbb{I}\Big\{\mathsf{T}:\ |\alpha_{\mathsf{T}}-\beta_{\mathsf{T}}|\sum_j \nu_{+,\mathsf{T}}\langle x_j\rangle a_{\phi,j}^{l,\mathsf{T}} \geq \eta \text{ eventually in } l\Big\},$$

and $\eta \mapsto \frac{\eta}{2+\eta}$ increasing in η, for every $\eta > 0$, except possibly countably many, such that $\eta \leq \eta_0(\varepsilon_0)$ and $\frac{\eta}{2+\eta} \leq \varepsilon_0/4$, from (4.29) and (4.31) we get that,

$$0 \geq \frac{1}{2}\lambda_\delta\Big(\frac{\eta}{2+\eta}\Big)^2\Big\{\mathbb{E}_\mu\Big[\int \beta_{\mathsf{T}}\liminf_{l\to\infty}\mathbb{I}(A_\phi^{\mathsf{T},l}=1)\mathfrak{m}_+(d\overline{\nu}_{+,\mathsf{T}})\Big] - \frac{\eta}{2+\eta}\Big\} \geq \frac{1}{8}\lambda_\delta\varepsilon_0\Big(\frac{\eta}{2+\eta}\Big)^2 \geq 0.$$

Therefore we arrive at a contradiction and hence proof is done. □

Proof of Theorem 1.7(b): As a first step we will show that $\overline{\nu}_+ = \mu\otimes\nu_{+,\mathsf{T}}$. Since for each $t > 0$, $\mathbb{P}_n^t \Rightarrow \overline{\nu}_+^t$, $\mathbb{P}_n^t(I_n) \Rightarrow \mathfrak{m}_+^t$, and $\mathbb{P}_n^t = \mathbb{E}_{U_n}(\mathbb{P}_n^t(I_n))$, the result follows by noting that $\nu_{+,\mathsf{T}}$ is an extremal measure.

Since μ is ergodic, from the proof of Lemma 4.4 it follows that $\rho^l_{-,\mathsf{T}} \to -\rho = -\mathbb{E}_\mu[\nu_{+,\mathsf{T}}\langle x_\phi\rangle]$ as $l \to \infty$. Now choosing $\eta < \rho/2$ for large l, $\rho^l_{-,\mathsf{T}} \le \eta$ and therefore following the proof of Lemma 4.5 (a) we obtain,

$$\nu_{+,\mathsf{T}}\Big(F^l_\phi = 1\Big) + \nu_{-,\mathsf{T}}\Big(F^l_\phi = 0\Big) \stackrel{\mathcal{P}}{\to} 0, \text{ as } l \to \infty.$$

Further using unimodularity of μ and DCT,

$$\nu_{+/-,\mathsf{T}}\Big[\sum_{i\in\partial\phi} \mathbb{I}(F^l_\phi \neq F^l_i)\Big] \stackrel{\mathcal{P}}{\to} 0, \text{ as } l \to \infty.$$

Now to complete the proof first choose an $\eta < \rho/2$, such that Lemma 4.1 and Lemma 4.5 holds, and then define

$$S_{\eta,l} = \{i : F^l_i = 1\}.$$

Now rest of the steps of the proof can be done by adapting the proof of Theorem 1.7(a). We omit the details. □

5. Continuity of $U(\cdot, 0)$ in β and edge-expander property

In this section we first show that under some assumptions on $\mu \in \mathcal{U}$, the function $\beta \mapsto U(\beta, 0)$ is continuous. Before going to the proofs we need to introduce few notations: For any infinite tree T, let $m^{+,t}_{\mathsf{T}} \equiv m^{+,t,\beta,B}_{\mathsf{T}}$ be the root magnetization for the Ising model with parameters β, B on $\mathsf{T}(t)$ with $+$ boundary condition and let $h^{+,t}_{\mathsf{T}} = \operatorname{atanh} m^{+,t}_{\mathsf{T}}$. Using Griffith's inequality, $\lim_{t\to\infty} h^{+,t}_{\mathsf{T}} = h^+_{\mathsf{T}}$ exists and we also get, for any positive integer t, any $i \in \partial\mathsf{T}(t)$,

$$h^+_{i\to\phi} = \sum_{i\to j} \operatorname{atanh}\Big(\tanh(\beta)\tanh(h^+_{j\to\phi})\Big), \tag{5.1}$$

where we write $v \to w$ to denote v is a parent of w, and $h^+_{x\to y} \equiv h^+_{\mathsf{T}_{x\to y}}$. Note that the above equality is also true for ϕ, i.e. the equality remains valid when LHS is replaced by h^+_{T} and in the RHS sum is taken over all $i \in \partial\phi$. For consistency of notations we will write $h^+_{\phi\to\phi} \equiv h^+_{\mathsf{T}}$.

When T is random, so is $h^+_{i\to\phi}$ and thus we get a collection of random variables $\{h^+_{i\to\phi}\}_{i\in V(\mathsf{T})}$.

The proof of $\beta \mapsto U(\beta, 0) \in \mathcal{C}$ is broken into two parts: (i) above criticality regime is proved in Lemma 5.3, and (ii) at criticality it is proved in Lemma 5.6. Thus Lemma 1.13 will follow from these two lemmas, once we show that assumptions of these two lemmas are satisfied for UGW and UMGW measures.

Proof of Lemma 1.13: Here we check that all the conditions required for Lemma 5.3 and Lemma 5.6 are satisfied:

• *Ergodicity:* To apply Lemma 5.3, and Lemma 5.6 we first need to show that the discrete time SRW on UMGW trees, conditioned on non-extinction, when the initial distribution biased by degree of the root, is ergodic.

To this end, let the measure biased by the degree of the root, be called as *augmented multi-type Galton-Watson* (AMGW) measure. Note that this measure can be defined as follows: choose an edge (v, w) of type (a, b) with probability $\pi(a, b) \propto \theta(a)A(a, b)$. Then construct two trees T_v and T_w, on the two sides of the edge (v, w), such that first generation of T_v are chosen with kernel $\rho_{a,b}$, that of T_w according to $\rho_{b,a}$, and from next generation onwards, in both the trees, children are chosen according to ρ. Now join T_v, and T_w, by the edge (v, w) and then fix the root of the final tree at v. It is easy to note that this a generalization of the AGW measure defined in [29, §3]. Now note that it is enough to prove that discrete time SRW on AMGW trees, conditioned non-extinction is ergodic.

By adapting the proof of Proposition 16.12, and the discussions after that in [30, pp. 558-559], and defining *regeneration point* corresponding to one particular edge-type, proof of ergodicity of SRW of an AMGW tree, conditioned on non-extinction can be completed. Details are omitted. Therefore, the same is proved any UMGW measure, conditioned on non-extinction.

• *Branching number.:* To apply Lemma 5.3 we need to show that, conditioned on non-extinction, for UMGW a.s. every tree T, for $i \in \partial\phi$, $\beta_c(\mathsf{T}) = \beta_c(\mathsf{T}_{i\to\phi})$, whenever $\mathsf{T}_{i\to\phi}$ is infinite. Note that it is enough to prove the same for AMGW trees, conditioned on non-extinction. Further noting that there is a one-one relation between $\operatorname{br}\mathsf{T}$ and $\beta_c(\mathsf{T})$, we would be done, if we prove the same relation with β_c replaced by branching number. Now for any tree T, let T^∞ be the set of vertices v with *infinite line of descent*. Then adapting the proof of [13, Appendix 6], it follows that conditional non-extinction T^∞ is a.s. an infinite tree without leaves, and follows a *modified* MGW distribution with a modified parameter (the off-spring distribution of the roots and that of the rest of the vertices could be different). Furthermore, conditioned on non-extinction, T is always the infinite tree T^∞ plus a bunch of finite trees emanating from each vertex of T^∞. Since branching number is not affected by finite sub-trees, we need to prove the result only for T^∞. Now noting that $\operatorname{br}\mathsf{T}^\infty = \max_{i\in\partial\phi}\operatorname{br}\mathsf{T}^\infty_{i\to\phi}$, T^∞, and conditioned on non-extinction being a modified MGW infinite tree, with no leaves, the result follows from [28, Proposition 6.5]: which says that for any MGW tree, for any type of the root-vertex, the branching number is almost surely the Perron-Frobenius eigenvalue of the mean matrix, on non-extinction.

• *W_T finite a.s.:* This condition is needed in Lemma 5.6. To this end, note that enough to show the condition holds for AMGW trees conditioned on non-extinction. From [28, Proposition 6.5], and from the arguments above it follows that, conditioned on non-extinction $\operatorname{br}\mathsf{T} = \operatorname{br}\mathsf{T}^\infty = \rho$, the Perron-Frobenius eigenvalue of the mean matrix, corresponding to the off-spring distribution of the children of the root in the modified MGW distribution. From [13, Appendix 6] it further follows that ρ is also the Perron-Frobenius eigenvalue of the mean matrix corresponding to the off-spring distribution of the children of the root of the original AMGW distribution. Now the result follows from [25, Theorem 1], since $\operatorname{br}\mathsf{T} = \rho$, on non-extinction.

Thus all the conditions of Lemma 5.3 and Lemma 5.6 are satisfied for UMGW measure. Since UGW is a special case of UMGW measure, all conditions are automatically satisfied for it. This completes the proof of Lemma 1.13. □

Lemma 5.1. *Consider any ergodic unimodular measure μ. For any T, let $\widetilde{m}^l(\mathsf{T})$ be the root magnetization of the Ising measure on the $\mathsf{T}(l)$, with $B_i = h^{+,0}_{i\to\phi}$ on for $i \in \partial\mathsf{T}(l)$, where $\{h^{+,0}_{i\to\phi}\}_{i\in V(\mathsf{T})}$ satisfies (5.1) with $\beta = \beta_0$. Let $m^+_l(\mathsf{T})$ be the same with $B_i = h^+_{i\to\phi}$ for $i \in \partial\mathsf{T}(l)$. Further assume that for every $\beta > 0$, and μ-a.e. T, for $i \in \partial\phi$, $\beta_c(\mathsf{T}) = \beta_c(\mathsf{T}_{i\to\phi})$, whenever $\mathsf{T}_{i\to\phi}$ is infinite.*
Then for any $\beta > \beta_0 > \beta_c$ and for any $l > 0$ integer,

$$m^l_+(\mathsf{T}) - \widetilde{m}^l(\mathsf{T}) \le M/l, \tag{5.2}$$

for μ-a.e. T, where M is a constant depending on β, β_0 and β_c.

There is another classical result, known as GHS inequality, (see [21]) which we have used in the proof below. This inequality is about the effect of the magnetic field $\underline{B}$ on local magnetizations at various vertices:

Proposition 5.2. [GHS inequality] *Let $\beta \ge 0$ and for $\underline{B} = \{B_i : i \in V\}$, denote by $m_j(\underline{B}) \equiv \mu\{\underline{x} : x_j = +1\} - \mu\{\underline{x} : x_j = -1\}$ the local magnetization at vertex j in the Ising model (3.2). If $B_i \ge 0$, for all $i \in V$, then for any three vertices $j, k, l \in V$ (not necessarily distinct),*

$$\frac{\partial^2 m_j(\underline{B})}{\partial B_k \partial B_l} \le 0.$$

Proof: For any positive integer k and a collection of positive numbers $\{H_i\}_{i\in\partial\mathsf{T}(k)}$ let

$$\mu^{k,\{H_i\}_{i\ \in\partial\mathsf{T}(k)}}(\underline{x}) \equiv \frac{1}{Z^{k,\{H_i\}_{i\in\partial\mathsf{T}(k)}}}\exp\Big\{\sum_{(i,j)\in\mathsf{T}(k)} x_ix_j + \sum_{i\in\partial\mathsf{T}(k)} H_ix_i\Big\} \tag{5.3}$$

and $m^k(\{H_i\}_{i\in\partial\mathsf{T}(k)})$ would be the root magnetization under the above Ising measure on $\mathsf{T}(k)$. For simplicity of notation we will often denote this root magnetization by $m^k(\{H_i\})$. Using Proposition 3.3 we note that $\widetilde{m}^{k+1} = m^k\Big(\Big\{\sum_{i\to j}\operatorname{atanh}\big(\tanh(\beta)\tanh(h^{0,+}_{j\to\phi})\big)\Big\}\Big)$. Now define the following function,

$$f(\tau) = \frac{\operatorname{atanh}(\tanh(\beta)\tau)}{\operatorname{atanh}(\tanh(\beta_0)\tau)},\ \tau\in(0,1]. \tag{5.4}$$

Noting that $\lim_{\tau\to 0} f(\tau) > 1$, the fact that any continuous function on a compact interval achieves its minima, we claim $\inf_{\tau\in(0,1]} f(\tau) > 1+\varepsilon$, for some $\varepsilon > 0$. Thus by Griffith's inequality,

$$\widetilde{m}^{k+1} \geq m^k\Big(\Big\{\sum_{i\to j}(1+\varepsilon)\operatorname{atanh}\big(\tanh(\beta_0)\tanh(h^{0,+}_{j\to\phi})\big)\Big\}\Big) = m^k(\{(1+\varepsilon)h^{0,+}_{i\to\phi}\}). \tag{5.5}$$

Since for μ-a.e. T, for $i\in\partial\phi$, $\beta_c(\mathsf{T}_{i\to\phi}) = \beta_c(\mathsf{T})$ whenver $\mathsf{T}_{i\to\phi}$ is infinite, by unimodularity of μ it follows that for μ-a.e. T, every $i\in V(\mathsf{T})$, and all $j\in\partial i$, $\beta_c(\mathsf{T}_{j\to i}) = \beta_c(\mathsf{T})$, whenever $\mathsf{T}_{j\to i}$ is infinite. Now μ being ergodic, it further follows that $\beta_c(\mathsf{T})$ is constant, equals β_c for μ-a.e. every T.

Now recall from [2, Theorem 1] it follows, for any fixed T, and any $\beta > \beta_c(\mathsf{T}) \equiv \beta_c$,

$$m^\beta_+(\mathsf{T}) \geq \text{const.}[(\beta-\beta_c)/\beta_c]^{1/2}.$$

Thus in our setting, whenever $\mathsf{T}_{i\to\phi}$ is infinite, $i\in V(\mathsf{T})$, $h^{0,+}_{i\to\phi}\geq B$ for some $B>0$. From (5.1) it further follows that, $h^{0,+}_{i\to\phi}\geq \xi\widetilde{\Delta}_i$, where $\xi\equiv\xi(\beta_0,B) = \operatorname{atanh}(\tanh(\beta_0)\tanh(B))$, and $\widetilde{\Delta}_i$ is the number of vertices j such that $j\sim i$ in $\mathsf{T}_{i\to\phi}$, and $\mathsf{T}_{j\to\phi}$ is infinite.

From (5.1) it further follows that $h^+_{i\to\phi}\leq\beta\widetilde{\Delta}_i$, for any $i\in V(\mathsf{T})$. Thus by Griffith's inequality,

$$m^{k+1}_+ \leq m^k(\{\beta\widetilde{\Delta}_i\}) \leq m^k(\{(\beta/\xi)h^{0,+}_{i\to\phi}\}). \tag{5.6}$$

By GHS inequality, $H\mapsto m^k(\{Hh^{0,+}_{i\to\phi}\})$ is concave and thus choosing ε small enough such that $1+\varepsilon<\beta/\xi$, and considering $H=\beta/\xi,\ (1+\varepsilon)$ and 1 we get,

$$m^k(\{(\beta/\xi)h^{0,+}_{i\to\phi}\}) - m^k(\{h^{0,+}_{i\to\phi}\}) \leq M\Big[m^k(\{(1+\varepsilon)h^{0,+}_{i\to\phi}\}) - m^k(\{h^{0,+}_{i\to\phi}\})\Big], \tag{5.7}$$

where $M=\frac{\beta-\xi}{\varepsilon\xi}$, ans thus combining (5.5), (5.6), and (5.7) we get,

$$m^{k+1}_+(\mathsf{T}) - \widetilde{m}^{k+1}(\mathsf{T}) \leq M[\widetilde{m}^{k+1}(\mathsf{T}) - \widetilde{m}^k(\mathsf{T})] \tag{5.8}$$

Since $m^k_+(\mathsf{T})$ is decreasing in k and $\widetilde{m}^k(\mathsf{T})$ increasing in k, summing the above inequality in k, we get

$$l[m^l_+(\mathsf{T}) - \widetilde{m}^l(\mathsf{T})] \leq \sum_{k=1}^{l}[m^k_+(\mathsf{T}) - \widetilde{m}^k(\mathsf{T})] \leq M\sum_{k=1}^{l}[\widetilde{m}^k(\mathsf{T}) - \widetilde{m}^{k-1}(\mathsf{T})] \leq M,$$

and hence the proof is complete. □

Lemma 5.3. *Under the assumptions of Lemma 5.1, $U(\beta,0)$ is continuous on (β_c,∞).*

Proof of Lemma 5.3: Right continuity of $U(\beta,0)$ follows from the proof of Lemma 3.1. Thus we only need to argue that it is left continuous. To this end for any $\beta \in (\beta_c,\infty)$ fix an $\beta_0 \in (\beta_c,\beta)$. Now for any infinte tree T and an integer $l \ge 1$, consider the Ising measure $\mu^{l,\{h^{0,+}_{i\to\phi}\}}$ on $\mathsf{T}(l)$, where $\{h^{0,+}_{i\to\phi}\}$ satisfy (5.1) $\beta=\beta_0$. Now define

$$U_l(\beta) = \frac{1}{2}\mathbb{E}_\mu\Big[\sum_{i\in\partial\phi} \mu^{l,\{h^{0,+}_{j\to\phi}\}}\langle x_0\cdot x_i\rangle\Big]. \tag{5.9}$$

First note that $U_l(\beta)$ are continuous in β and by Griffith's inequality $U_l(\beta)$ increasing in β. Now by Proposition 3.3, we also have

$$U_{l+1}(\beta) = \frac{1}{2}\mathbb{E}_\mu\Big[\sum_{i\in\partial\phi} \mu^{l,\{\sum_{k\to j}\operatorname{atanh}[\tanh(\beta)\tanh(h^{0,+}_{j\to\phi})]\}}\langle x_0\cdot x_i\rangle\Big]. \tag{5.10}$$

Since $\beta>\beta_0$,

$$h^{0,+}_{k\to\phi} \le \sum_{k\to j}\operatorname{atanh}[\tanh(\beta)\tanh(h^{0,+}_{j\to\phi})].$$

Thus by Griffith's inequality we conclude that $U_l(\beta)$ are increasing in l also. Once we further show that $U_l(\beta)\uparrow U(\beta,0)$, interchanging the limit in l and β, we will obtain $U(\beta,0)$ is left continuous in β and proof will be done.

To prove $U_l(\beta)\uparrow U(\beta,0)$ we first note that repeating the same arguments as in Lemma 5.1 we get,

$$\sum_{i\in\partial\phi}\Big(m^l_{+,i\to\phi}(\mathsf{T}) - \widetilde{m}^l_{i\to\phi}(\mathsf{T})\Big) \le \frac{M\Delta_\phi}{l-1},$$

where for any T, $m^l_{+,i\to\phi}(\mathsf{T})$ (and $\widetilde{m}_{i\to\phi}(\mathsf{T})$) is the root magnetization of $T_{i\to\phi}$ under the Ising measure on $\mathsf{T}(l)$ with $\{h^{+}_{j\to\phi}\}$ boundary condition (and with $\{h^{0,+}_{j\to\phi}\}$ boundary condition, respectively). Now repeating the same steps as in [10, Theorem 4.2], we have,

$$||\mu^l_{+,\mathsf{T}(2)} - \widetilde{\mu}^l_{\mathsf{T}(2)}||_{TV} \le C\frac{\Delta_\phi}{l-1}, \tag{5.11}$$

for some constant C. Here $\mu^l_{+,\mathsf{T}(2)}$ (and $\widetilde{\mu}_{\mathsf{T}(2)}$) is the marginal measure on $\mathsf{T}(2)$ induced by the Ising measure on $\mathsf{T}(l)$ with $\{h^{+}_{j\to\phi}\}$ boundary condition (and with $\{h^{0,+}_{j\to\phi}\}$ boundary condition respectively). Thus for $f=\frac{1}{\Delta_\phi}\sum_{i\in\partial\phi}x_\phi x_i$,

$$0 \le \lim_{l\to\infty}(\mu^l_{+,\mathsf{T}(2)}(f) - \widetilde{\mu}^l_{\mathsf{T}(2)}(f)) \le C\lim_{l\to\infty}\frac{\Delta_\phi}{l-1} = 0. \tag{5.12}$$

Now the result follows by noting that $\mathbb{E}_\mu[\Delta_\phi]<\infty$ and a use of DCT. $\square$

Remark 5.4. Proof of Lemma 1.15 can be done by imitating the proof of [10, Lemma 2.3]. We omit the details. In [10] many results are proved based on [10, Lemma 2.3, Lemma 4.3]. Here also we have obtained similar results in Lemma 5.1 and Lemma 1.15 and thus as a consequence many similar results as in [10] could also be obtained here.

In Lemma 5.3 we have already shown that for $\beta\in(\beta_c,\infty)$, $U(\beta,0)$ is continuous in β under some appropriate assumptions. Now for $\beta<\beta_c$ there is only one Gibbs measure on any tree T and thus $U(\beta,0)$ is also continuous for any $\beta<\beta_c$. Now we show that for $U(\beta,0)$ is continuous also at $\beta=\beta_c$ under some appropriate assumptions.

To prove this result we will use some results from [36]. Before stating the results we first need to discuss few definitions and notations:

Let T be any tree rooted at o. For finite trees $\widetilde{\partial}\mathsf{T}$ will denote the set of all leaves different from o and for any infinite tree $\widetilde{\partial}\mathsf{T}$ will denote the set of all infinite non-backtracking paths emanating from root (termed as *rays*). Note that for any finite tree there is a one-one correspondence between the leaves of the trees and the non-backtracking paths from root to leaves. Thus for any finite tree T, when we write $y \in \widetilde{\partial}\mathsf{T}$ we will often identify y with the path from o to y.
Let $\{R(e) : e \in E(\mathsf{T})\}$ be a collection of *resistances* (non-negative numbers) assigned to the edges of T. Let σ be a flow on T, i.e. σ is a non-negative function defined on $E(\mathsf{T})$ such that whenever $v \to w$ (v is parent of w) and w is not a leaf, $\sigma(vw) = \sum_{y:w\to y} \sigma(wy)$. Define, $|\sigma| := \sum_{y:o\to y} \sigma(oy)$. For $y \in \widetilde{\partial}\mathsf{T}$ define,

$$\begin{aligned}
V_\sigma(y) &:= \sum_{e\in y} (\sigma(e)R(e))^2, \\
V(\sigma) &:= \sup\{V_\sigma(y) : y \in \widetilde{\partial}\mathsf{T}\}, \\
\mathrm{cap}_3(\mathsf{T}) &:= \sup\{|\sigma|, \mu \text{ a flow on } \mathsf{T} \text{ with } V(\sigma) = 1\}.
\end{aligned}$$

Now we have the following proposition:

Proposition 5.5. [36, Theorem 3.2] *Let T by any finite tree rooted at o.Suppose there exists $\kappa_1 > 0$ and a collection of positive constants $\{a_v : v \in V(\mathsf{T})\}$ such that for every $v \in v(\mathsf{T})$ and $\xi \geq 0$,*

$$f_v(\xi) \leq \frac{a_v \xi}{(1 + (\kappa_1 \xi)^2)^{1/2}}. \tag{5.13}$$

Then any solution to the system

$$\xi_v \leq \sum_{v\to w} f_w(\xi_w) \text{ with } \xi_w = \infty \text{ when } w \text{ is a leaf,} \tag{5.14}$$

satisfies $\xi_o \leq \frac{\mathrm{cap}_3(\mathsf{T})}{\kappa_1}$ with the resistances $R_v = \prod_{0\leq y\leq v} a_y^{-1}$, to the edge between v and its parent, where $\prod_{0\leq y\leq v}$ denote the product over all indices y that are present on the path between o and v.

For any infinite tree T and any positive integer N if we consider $\mathsf{T}(N)$, then in [36, Lemma 4.2] the authors show that $\{\xi_v,\ v \in V(\mathsf{T}(N))\}$ satisfy the system of equations (5.14) for some appropriate choices of f_v's with $a_v = \tanh(\beta)$, and for any $v \in V(\mathsf{T}(N))$, ξ_v is the log-likelihood ratio of having spin 1 versus -1 at v, given the plus boundary condition on $\mathsf{T}(N)$. In [36, Lemma 4.3] it is also shown that f_v's satisfy (5.13) and thus by Proposition 5.5 for any infinite tree T and every positive integer N, $\xi_o^{(N)} \leq \frac{\mathrm{cap}_3(\mathsf{T}(N))}{\kappa_1}$, where $\xi_o^{(N)}$ is the log-likelihood ratio of having spin 1 versus -1 with plus boundary condition on the N-th level. If we can show that $\mathrm{cap}_3(\mathsf{T}(N)) \stackrel{a.s.}{\to} 0$ as $N \to \infty$ then combining the results we would obtain

$$\xi_o := \log\left[\frac{\nu_{+,\mathsf{T}}(x_\phi = +1)}{\nu_{+,\mathsf{T}}(x_\phi = -1)}\right] = \lim_{N\to\infty} \xi_o^{(N)} = 0, \text{ a.s.}, \tag{5.15}$$

and thus there would be only one Ising-Gibbs measure and hence $U(\beta, 0)$ would be continuous at $\beta = \beta_c$. Since T might have leaves, showing $\mathrm{cap}_3(\mathsf{T}(N)) \to 0$ is a bit difficult. Instead we proceed via the following approach:
For any tree T, whenever we come across a leaf in T we construct a *ray* from that leaf to infinity. Let $\widetilde{\mathsf{T}}$ denote the modified tree with extra rays. Thus $\widetilde{\partial}\widetilde{\mathsf{T}}(N) = \partial\widetilde{\mathsf{T}}(N)$ and by Griffith's inequality $\xi_o^{(N)} \leq \widetilde{\xi}_o^{(N)}$, where $\widetilde{\xi}_o^{(N)}$ is the corresponding quantity defined for $\widetilde{\mathsf{T}}$. Therefore we will be done if we can show that $\mathrm{cap}_3(\widetilde{\mathsf{T}}(N)) \stackrel{a.s.}{\to} 0$.

Lemma 5.6. *For any ergodic unimodular measure on trees, such that* $W_{\mathsf{T}} := \limsup_{k\to\infty} \frac{|\partial \mathsf{T}(k)|}{(\mathrm{br}\,\mathsf{T})^k} < \infty$, *for a.s. every* T, *the function* $\beta \mapsto U(\beta,0)$ *is continuous at* $\beta=\beta_c$.

Proof: Fix any positive integer N and consider any flow σ on $\widetilde{\mathsf{T}}(N)$ such that $|\sigma|=1$. Then for any probability measure on $\{y : y \in \partial\widetilde{\mathsf{T}}(N)\}$,

$$V(\sigma) \ge \sum_{y\in\partial\widetilde{\mathsf{T}}(N)} \Big[\sum_{e\in y}\sigma^2(e)\tanh(\beta_c)^{-2|e|}\Big]p(y) = \sum_{k=1}^{N}\tanh(\beta_c)^{-2k}\sum_{|e|=k}\sigma^2(e)\sum_{y\ni e}p(y). \tag{5.16}$$

By a slight abuse of notation we let $p(e)=\sum_{y\ni e}p(y)$. Note that $p(e)$ thus defined for all edges e constitute a flow, $|p|=1$ and also $\sum_{|e|=k}p(e)=1$ for all k, since there no leaves in $\widetilde{\mathsf{T}}(N)$ except at the N-th generation. Now applying Cauchy-Schwartz inequality and taking $\sigma=p$,

$$\Big[\sum_{|e|=k}\sigma^2(e)p(e)\Big] \ge \Big(\sum_{|e|=k}\sigma(e)p(e)\Big)^2 = \Big(\sum_{|e|=k}\sigma^2(e)\Big)^2. \tag{5.17}$$

Using Cauchy-Schwartz inequality once again,

$$\Big(\sum_{|e|=k}\sigma^2(e)\Big) \ge \frac{1}{|\partial\widetilde{\mathsf{T}}(k)|}\Big(\sum_{|e|=k}\sigma(e)\Big)^2 = \frac{1}{|\partial\widetilde{\mathsf{T}}(k)|}. \tag{5.18}$$

Now from [27, Theorem 1.1] we know that $\mathrm{br}\,\mathsf{T}[\tanh(\beta_c)]=1$, thus from (5.16), (5.17) and (5.18), for any flow σ on $\widetilde{\mathsf{T}}(N)$, with $|\mu|=1$,

$$V(\mu) \ge \sum_{k=1}^{N}\left[\frac{(\mathrm{br}\,\mathsf{T})^k}{|\partial\widetilde{\mathsf{T}}(k)|}\right]^2 \tag{5.19}$$

Note that by construction, $|\partial\widetilde{\mathsf{T}}(k)| \le |\partial\mathsf{T}(k)| + |\partial\widetilde{\mathsf{T}}(k-1)|$. Hence,

$$\limsup_{k\to\infty}\frac{|\partial\widetilde{\mathsf{T}}(k)|}{(\mathrm{br}\,\mathsf{T})^k} \le W_{\mathsf{T}} + (\mathrm{br}\,\mathsf{T})^{-1}\limsup_{k\to\infty}\frac{|\partial\widetilde{\mathsf{T}}(k)|}{(\mathrm{br}\,\mathsf{T})^k}$$

$$\text{i.e.}\quad \limsup_{k\to\infty}\frac{|\partial\widetilde{\mathsf{T}}(k)|}{(\mathrm{br}\,\mathsf{T})^k} \le \frac{\mathrm{br}\,\mathsf{T}}{\mathrm{br}\,\mathsf{T}-1}W_{\mathsf{T}},$$

and thus

$$\sum_{k=1}^{N}\left[\frac{(\mathrm{br}\,\mathsf{T})^k}{|\partial\widetilde{\mathsf{T}}(k)|}\right]^2 \stackrel{a.s.}{\to} \infty \text{ as } N\to\infty,$$

and hence the proof follows by the discussion above. $\square$

Lemma 1.17 is well known. However we provide its proof for completeness.

Proof of Lemma 1.17: For $i\in\mathcal{Q}$, and $\underline{k}\in\mathbb{Z}_+^{\mathcal{Q}}$, let

$$\alpha_{i,\underline{k}} = \theta(i)P_i(\underline{k}), \text{ and } \boldsymbol{\alpha} = (\alpha_{i,\underline{k}})_{i\in\mathcal{Q},\underline{k}\in\mathbb{Z}_+^{\mathcal{Q}}}.$$

From Definition 1.10, it follows that there are $n\alpha_{i,\underline{k}}(1+o(1))$ vertices of type $i\in cQ$, and neighborhood configuration $\underline{k}$. Fix any $\delta_0\le 1/2$, and consider any vector $\boldsymbol{\delta}=(\delta_{i,\underline{k}})_{i\in\mathcal{Q},\underline{k}\in\mathbb{Z}_+^{\mathcal{Q}}}$, such that $|\boldsymbol{\delta}|\in(\delta_0,1/2)$. Let $U^{\boldsymbol{\delta}}$ denote a subset of vertices of size $n\boldsymbol{\delta}(1+o(1))$, among which $n\delta_{i,\underline{k}}(1+o(1))$ are of type i and neighborhood configuration $\underline{k}$. Let $S_{\boldsymbol{\delta}}^{\varepsilon}$ be the set of configurations for which there are exists a set $U^{\boldsymbol{\delta}}$ with $n\varepsilon(1+o(1))$ edges with one end in $U^{\boldsymbol{\delta}}$ and the other in $U_{\boldsymbol{\delta}}^c$.

It is enough to show that for every $\boldsymbol{\delta} \in (\delta_0, 1/2)$ there exists an $\varepsilon := \varepsilon(\delta_0) > 0$ such that

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}(S^{\varepsilon}_{\boldsymbol{\delta}}) < 0, \tag{5.20}$$

uniformly overall all choices of $\boldsymbol{\delta}$. To this end we first note that

$$\begin{aligned} \frac{1}{n}\log \mathbb{P}(S^{\varepsilon}_{\boldsymbol{\delta}}) &= \frac{1}{n}\log \#\Big\{\text{choices possible for } U^{\boldsymbol{\delta}}\Big\} + \frac{1}{n}\log \mathbb{P}\Big\{ \text{ such choice matches with itself}\Big\} \\ &:= N_{\boldsymbol{\delta}} + Q_{\boldsymbol{\delta}}. \end{aligned}$$

For ease of notations further define, $\hat{\boldsymbol{\alpha}} := (\hat{\alpha}_{i,j})_{i,j\in\mathcal{Q}}$, where for every every $i,j \in \mathcal{Q}$, $\hat{\alpha}_{i,j} = \sum_{\underline{k}} k_j \alpha_{i,\underline{k}}$, and similarly define $\hat{\boldsymbol{\delta}}$. Now using the approximations,

$$\frac{1}{n}\log n! = \log\big(\frac{n}{e}\big) + o(1), \text{ and } \frac{1}{n}\log n!! = \frac{1}{2}\log\big(\frac{n}{e}\big) + o(1),$$

we get

$$N_{\boldsymbol{\delta}} \approx \sum_{i,\underline{k}} \alpha_{i,\underline{k}} H\Big(\frac{\delta_{i,\underline{k}}}{\alpha_{i,\underline{k}}}\Big) = \sum_i \sum_j \sum_{\underline{k}} \frac{k_j \alpha_{i,\underline{k}}}{|\underline{k}|} H\Big(\frac{\delta_{i,\underline{k}}}{\alpha_{i,\underline{k}}}\Big)$$

and

$$Q_{\boldsymbol{\delta}} \approx -\frac{1}{2}\sum_{i,j\in\mathcal{Q}} \hat{\alpha}_{i,j} H\Big(\frac{\hat{\delta}_{i,j}}{\hat{\alpha}_{i,j}}\Big),$$

where $H(p) = -p\log p - (1-p)\log(1-p)$, $p \in [0,1]$. By the concavity of $H(\cdot)$, and noting that $|\underline{k}| \ge 3$, we have for any $i,j \in \mathcal{Q}$ that

$$\sum_{\underline{k}} \frac{k_j\alpha_{i,\underline{k}}}{|\underline{k}|} H\Big(\frac{\delta_{i,\underline{k}}}{\alpha_{i,\underline{k}}}\Big) - \frac{1}{2}\hat{\alpha}_{i,j} H\Big(\frac{\hat{\delta}_{i,j}}{\hat{\alpha}_{i,j}}\Big) \le \frac{1}{3}\sum_{\underline{k}} k_j\alpha_{i,\underline{k}} H\Big(\frac{\delta_{i,\underline{k}}}{\alpha_{i,\underline{k}}}\Big) - \frac{1}{2}\hat{\alpha}_{i,j} H\Big(\frac{\hat{\delta}_{i,j}}{\hat{\alpha}_{i,j}}\Big) \le -\frac{1}{6}\hat{\alpha}_{i,j} H\Big(\frac{\hat{\delta}_{i,j}}{\hat{\alpha}_{i,j}}\Big).$$

Summing the above for $i,j \in \mathcal{Q}$, and using concavity again, we obtain an upper bound, less than 0, depending only on $|\boldsymbol{\delta}|$. Now the compactness of the set of choices of $\boldsymbol{\delta}$ yields (5.20) for $\varepsilon = 0$. Similar calculations can be done for any $\varepsilon = 0$. For example when $|\mathcal{Q}| = 1$,

$$\limsup_{n\to\infty} \frac{1}{n}\log \mathbb{P}(S^{\varepsilon}_{\boldsymbol{\delta}}) \le \frac{\hat{\alpha}}{6} H\Big(\frac{\hat{\delta}}{\hat{\alpha}}\Big) + \frac{\hat{\delta}}{2} H\Big(\frac{\varepsilon}{\hat{\delta}}\Big) + \frac{1}{2}(\hat{\alpha} - \hat{\delta}) H\Big(\frac{\varepsilon}{\hat{\alpha} - \hat{\delta}}\Big).$$

Since the bound above is continuous in ε, there exists a ε such that (5.20) holds. For $|\mathcal{Q}| > 1$ computations are similar. We omit the details. □

References

[1] M. Aizeman. Translation invariance and instability of phase coexistence in the two-dimensional Ising system. *Comm. Math. Phys.*, 73, 8394, 1980.

[2] M. Aizenman, D. Barsky, and R. Fernández. The phase transition in a general calss of Ising-type models is sharp. *J. State. Phys.* 47, 343-374, 1987.

[3] M. Aizenmen, J. Wehr. Rounding of first-order phase transitions in systems with quenched disorder. *Communications in Mathematical Physics*, 130 489-530, 1990.

[4] D. Aldous and R. Lyons. Processes on unimodular random networks. *Electron. J. Probab.*, 12(54):1454–1508, 2007.

[5] D. Aldous and J. Steele. The objective method:probabilistic combinatorial optimization and local weak convergence. in H. Kesten, editor, *Probability on Discrete Structures,* volume 110 of *Encyclopedia Math Sci.,* pages 1-72. Springer, Berlin. Probability Theory, 1.

[6] K. B. Athreya, and P. E. Ney. *Branching Process.* Dover Publications Inc., Mineola, NY, 2004. Reprint of the 1972 original [Springer, New York; MR0373040].

[7] I. Benjamini and O. Schramm. Recurrence of distributional limits of finite planar graphs. *Electron. J. Probab.*, 6(23):13 pp. (electronic), 2001.

[8] N. Berger, C. Kenyon,E. Mossel,and Y. Peres. Glauber dynamics on Trees and Hyperbolic graphs. *Probab. Theory Related Fields*, Vol. 131 No. 3, 311-340, 2005.

[9] T. Bodineau. Translation invariant Gibbs states for the Ising model. *Probability Theory and Related Fields*, 135, 153-186, 2006.

[10] A. Dembo, and A. Montanari. Ising models on locally tree-like graphs. *Ann. Appl. Probab.*, 20(2):565–592, 2010.

[11] A. Dembo, and A. Montanari. Gibbs measures and phase transitions on random graphs. *Brazilian Journal of Probability and Statistics*, 24(2): 137-211, 2010.

[12] A. Dembo, A. Montanari, and N. Sun. Factor models on locally tree like graphs. *preprint*, 2011.

[13] A. Dembo, and N. Sun. Central limit theorem for biased random walk on Galton-Watson trees. *Elect. J. Probab.*, (to appear), 2011.

[14] R. Dobrushin, and S. Shlosman. The problem of translation invariance of Gibbs states at low temperatures. *Mathematical physics reviews*, 5, 53195, 1985.

[15] S. Dommers, C. Giardinà, and R. van der Hofstad. Ising models on power-law random graphs. *J. Stat. Phys.*, 141(4):638–660, 2010.

[16] F. J. Dyson. Existence of a phase-transition in a one-dimensional Ising ferromagnet. *Comm. Math. Phys.*, 12(2) :91-107, 1969.

[17] R. S. Ellis, and C. M. Newman. The statistics of Curie-Weiss models. *Journal of Statistical Physics*, 19, 149-161, 1978.

[18] H. Georgii. *Gibbs measures and phase transitions*, volume 9 of *de Gruyter Studies in Mathematics*. Walter de Gruyter & Co., Berlin, 1988.

[19] H. O. Georgii, and Y. Higuchi. Percolation and number of phases in the two-dimensional Ising model. *Journal of Mathematical Physics*, 41, 11531169, 2000.

[20] A. Greschenfield, and A. Montanari. Reconstruction for models on random graphs. *48nd Annual Symposiums of Foundations of Computer Science*, Providence, RI, 2007.

[21] R. B. Griffiths, C. A. Hurst, and S. Sherman. Concavity of magnetization of an Ising ferromagnet in a positive external field. *J. Math. Phys.*, 11, 790-795, 1970.

[22] G. Grimmett. *The random-cluster model*, volume 333 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 2006.

[23] O. Kallenberg. *Foundations of modern probability.* 2nd edition, Springer Series in Statistics, 2002.

[24] C. Külske. Metastates in disordered mean-field models: random field and Hopfield models. *Journal of Statistical Physics*, 88, 1257-1293, 1996.

[25] T. G. Kurtz, R. Lyons, R. Pemantle, and Y. Peres. A conceptual proof of the Kesten-Stigmum theorem for multitype branching processes. *Classical and Modern Branching Processes,* Springer, New York, 1997, pp. 181-185.

[26] T. M. Liggett. *Interacting particle systems*. Classics in Mathematics. Springer-Verlag, Berlin, 2005. Reprint of the 1985 original.

[27] R. Lyons. The Ising model and percolation on trees and tree-like graphs. Commum. Math. Phys., 125(2), 337-353, 1989.

[28] R. Lyons. Random walk and percolation on trees. *Ann. Probab.*, 18(3), 931-958, 1990.

[29] R. Lyons, R. Pemantle, and Y. Peres. Ergodic theory on Galton-Watson trees: speed of random walk and dimension of harmonic measure. *Ergodic Theory Dynam. Systems*, 15, 593-619, 1995.

[30] R. Lyons, and Y. Peres. *Probability on Trees and Networks.* Preprint.

[31] M. Mézard, and A. Montanari. *Infomation, physics, and computation*. Oxford Graduate Texts. Oxford University Press, Oxford, 2009.

[32] A. Montanari, E. Mossel, and A. Sly. The weak limit of Ising models on locally tree-like graphs. *Probab. Theory Related Fields*, 2011 (in press).

[33] M. E. J. Newmann. The structure and function of complex networks. *SIAM Review*, 45(2): 167-256, 2003.

[34] C. M. Newman, and D. L. Stein. Spatial inhomogeneity and thermodynamic chaos. *Physical Review letters*, 76, 4821-4824, 1996.

[35] M. Niss. History of Lenz-Ising model 1920-1950: from ferromagnetic to cooperative phenomena. *Archive for History of Exact Sciences*, 59(3): 267-318, 2005.

[36] R. Pemantle, and Y. Peres. The crictical Ising model on trees, concave recursions and nonlinear capacity. *Ann. Probab.*, 38, 184-206, 2010.

[37] L. De Sanctis, and F. Guerra. Mean field dilute ferromagnet: high temperature and zero temperature behavior.*Journal of Statistical Physics*, 132: 759-785, 2008.
[38] D. W. Stroock. Probability Theory, an analytic view. *Cambridge University Press*, 1994.

†Department of Mathematics, Stanford University
Building 380, Sloan Hall, Stanford, California 94305

*†Department of Statistics, Stanford University
Sequoia Hall, 390 Serra Mall, Stanford, California 94305